Denver Law Review

Volume 87 | Issue 1 Article 9

December 2020

Vol. 87, no. 1: Full Issue

Denver University Law Review

Published by the
University of Denver
Sturm College of Law

2009 Volume 87 Issue 1

DENVER UNIVERSITY LAW REVIEW

CONTENTS

IN MEMORIAM: ERIK B. BLUEMEL, AND THE LIFE OF THE LEGAL MIND

The place is here. The time is now.

– Erik B. Bluemel (1977–2009)

Erik Bluemel joined the faculty of the Sturm College of Law in the fall of 2008. His appointment as a professor of law was a dream come true for him, and he was comfortable enough to announce this triumphantly nearly every week. He was, as he repeatedly told me, "living the dream, baby." Tragically, Erik passed away just days after his second semester of teaching was complete. This issue of the *Denver University Law Review* is dedicated to the memory of Erik—his selfless service to others, his dry and witty humor, his wonderful compassion, and his inspiring intellect. Erik has left behind a large circle of friends at the Sturm College of Law and elsewhere, and we do ourselves a service by reflecting on his life.

I.

Although only thirty-one years old when he died, Erik lived a rich and accomplished life. He studied and distinguished himself at Berkeley, the University of Chile, New York University School of Law, and Georgetown University Law Center. He had backpacked throughout the world, studying cultures and, I imagine, developing what I came to love best about him, his *joie de vivre*. (Perhaps related, he was a masterful pool player and skilled card player.) He also gave presentations on topics of law and culture, even before he went to law school; indeed, Erik's expertise in global climate change had made him an emerging international commodity on the lecture circuit—he was scheduled to present in Montana and Italy in the weeks following his untimely death. He had already presented in Norway earlier in the year, and China before that. Erik particularly relished his trips abroad, where he could explore discussions with colleagues worldwide about the relationship between the complexity of government structure and potential future solutions to the problem of climate change.

Erik also had all the traits of a skilled and seasoned lawyer. I would consult with Erik on criminal cases—a field he knew relatively little about—because of his keen sense of adversarial strategy and awareness of the realities of litigation. Perhaps most impressive of all, given his age and other commitments and accomplishments, he was a prolific writer. Erik wrote well, he wrote often (he published 15 papers in less than 10 years), and his writing was important.

The loss of Erik leaves those of us who knew him feeling deprived on a personal and emotional level. For the scholarly community, his loss creates a void in the academic literature and activist movements relating to environmental and indigenous rights issues. In one early effort at what will be a long process of honoring Erik's memory and furthering his research and social agenda, the Environmental Law Clinic at the University of Denver has created the Erik B. Bluemel International Environmental Law Scholarship. Each year, a graduate level law student will be selected to work with an international environmental organization and study such issues as the relationship between environmental degradation and human rights.

II.

Erik brought youthful intellectual pride and spirit to the Sturm College of Law. He was eager to contribute to the law school community in ways that were both seen and unseen. He was a regular at the "Third Thursday" social events held in the forum of the law school, and not just because he liked free beer. Like me, Erik was in his first year of teaching, and he was eager to mingle with his new colleagues and students. Midway through his first semester at the law school, he had already compiled a list of the entire active faculty, whether he had talked to them personally, and what he had heard second-hand about them. To be sure, Erik's list was a source of a good many laughs between the two of us as we swapped stories about, for example, colleagues confusing us with students, but for Erik the list also reflected his genuine desire to know the DU community that he had joined. He was determined to become actively involved in the daily life of the students, staff, and faculty, and I think that it is fair to say that his enthusiasm served as a motivation for the rest of us to redouble our sense of pride and commitment in the University of Denver.

Indicative of the high esteem in which we as a faculty held Erik, in just his second semester of teaching he was elected to the Faculty Executive Committee tasked with many important issues of faculty governance, including the development of a strategic plan for the law school. My senior colleagues have mentioned to me that they saw qualities in Erik that clearly foreshadowed a future as one of the law school's rising leaders. As the Sturm College of Law works through an exciting transition period—hiring a permanent Dean, developing a long-term strategic plan for the school, reducing class size, and increasing the number of faculty—the DU community will do well to consider Erik's example. A tireless scholar and selfless contributor to our academic community, Erik also had exacting standards of excellence and took very personally the importance of challenging others—students, colleagues, and future colleagues—with the hard questions necessary to bring out the best in them.

III.

In addition to his remarkable record of extraordinary scholarship and service accomplishments at such an early age, Erik was quite simply a good human being. And to those he knew well (including this author), he was a good friend. Erik's candor, compassion, and hilarity were integral parts of my first year of teaching. There was not a moment of junior faculty stress or disappointment that, between the two of us in our common hallway, would not ultimately lead to eruptive and curing laughter. Like many of us on the faculty, the sense of loss is deep; nearly every law school debate leaves me wishing he were here to contribute, or at least to laugh about it afterwards. But I like to think that Erik has taught us all a little more about what it means to be an academic.

Simply put, Erik embodied the life of the mind; he was the consummate free and deep thinker. Early in Hannah Arendt's meditations on *The Life of The Mind*, she poses the following: "Could the activity of thinking as such, the habit of examining whatever happens to come to pass or to attract attention, regardless of results and specific content, could this activity be among the conditions that make men abstain from evil-doing or even actually 'condition' them against it?" Erik had a profound sense of integrity and empathy, and after each of our countless debates about every piece of trivia known in law, sports, entertainment, and vegetarianism (many of which ended with him winking, as silent acknowledgement that he had, at least partially, persuaded me), I would reflect on the beauty of his mind, or as I would often put it, his "huge brain." I continue to wonder whether he was such a great thinker because of his conscience and his desire to be good to all animals (human and non-human) and the environment, or whether he was such a good person because of his rich and curious intellect. I will never know for sure, but I know that he was a person of courageous integrity and profound intellect. He lived the life of the mind.

Justin Marceau[*]

* Assistant Professor, University of Denver Sturm College of Law.

AGAINST THE RIGHT TO BODILY INTEGRITY: OF CYBORGS AND HUMAN RIGHTS

GOWRI RAMACHANDRAN†

INTRODUCTION

Creating a list of fundamental human rights is a controversial project, but there is one right that appears in many lists—a right to bodily integrity, security, or control over one's own body.[1] The content of what the right should be is hotly contested.[2] For instance, does the right to bodily integrity require that organ selling be forbidden?[3] Or, do fundamental rights require that organ selling be permitted?[4] What about sales

---

† B.A., Yale College; M.A., Harvard University; J.D., Yale Law School. Visiting Associate Professor, University of California, Berkeley School of Law, Associate Professor of Law, Southwestern Law School. I have received helpful comments from Anshul Amar, Ronald Aronovsky, Molly Beutz, Devon Carbado, David Fagundes, Bryant Garth, John Greenman, Dante Harper, Sonia Katyal, Janine Kim, Sung-Hui Kim, Ethan Leib, Stephen Munzer, Camille Gear Rich, Angela Riley, Michael Scott, Patrick Shin, Dean Spade, Michael Waterstone, and participants at the Law and Society conference, the Southeastern Association of Law Schools conference, and the Lavender Law conference. The paper also benefited from my participation as a fellow at the Cardozo Law School Patenting People Conference. Mark Berkebile, Sylvia Chiu, Brian Craggs, Heather Croft, Michael Manapol, and John Trani provided helpful research assistance. The editors of the Denver University Law Review provided excellent editorial advice. Any errors are my own.

1. *See, e.g.*, Washington v. Glucksberg, 521 U.S. 702, 720 (1997) (citing Rochin v. California, 342 U.S. 165 (1952) (recognizing a right to "bodily integrity"); Universal Declaration of Human Rights, G.A. Res. 217A, at 71, 72, U.N. GAOR, 3d Sess., 1st plen. mtg., U.N. Doc. A/810 (Dec. 10, 1948), *available at* http://daccessdds.un.org/doc/RESOLUTION/GEN/NR0/043/88/IMG/NR004388.pdf?OpenElement (stating in Article 3 that "[e]veryone has the right to life, liberty and security of person"); MARTHA C. NUSSBAUM, WOMEN AND HUMAN DEVELOPMENT: THE CAPABILITIES APPROACH 78 (2000) (listing bodily integrity as one of ten "central human functional capabilities").

2. *See, e.g.*, Stephen R. Munzer, *An Uneasy Case Against Property Rights in Body Parts*, *in* PROPERTY RIGHTS 259, 261 (Ellen Frankel Paul et al. eds., 1994); Radhika Rao, *Property, Privacy, and the Human Body*, 80 B.U. L. REV. 359, 360–63 (2000) (describing the American constitutional right to bodily integrity and the fact that it has the character of a privacy interest—not a property interest—in that it protects against state interference, but does not protect a right to alienate the interest); *see also* Margo A. Bagley, *Patent First, Ask Questions Later: Morality and Biotechnology in Patent Law*, 45 WM. & MARY L. REV. 469, 511–12, 546 (2003) ("[T]he result may soon be, among other things, patents on human fetuses that are genetically modified in ways one can only imagine. Patent protection could convert such fetuses, to the extent they are denied constitutional protection, into justifiable commodities, supplying life-saving tissue and organs to sick children and adults.").

3. *See, e.g.*, U.S. DEP'T OF HEALTH & HUMAN SERVS., ORGAN TRANSPLANTATION: ISSUES AND RECOMMENDATIONS 99 (1986); HASTINGS CTR., ETHICAL, LEGAL AND POLICY ISSUES PERTAINING TO SOLID ORGAN PROCUREMENT: A REPORT OF THE PROJECT ON ORGAN TRANSPLANTATION 3–4 (1985); Munzer, *supra* note 2, at 266 (describing the Kantian human dignity argument as applied to organ sales); *see also* Susan M. Shell, *Kant's Concept of Human Dignity as a Resource for Bioethics*, *in* HUMAN DIGNITY AND BIOETHICS: ESSAYS COMMISSIONED BY THE PRESIDENT'S COUNCIL ON BIOETHICS 333, 344–45 (2008), *available at* http://www.bioethics.gov/reports/human_dignity/human_dignity_and_bioethics.pdf.

4. *Cf.* Eugene Volokh, *Medical Self-Defense, Prohibited Experimental Therapies, and Payment for Organs*, 120 HARV. L. REV. 1813, 1835 (2007). Volokh based his argument not in a right to "bodily integrity," but in a right to self-defense. *Id.* at 1815–16.

of the use of one's body, such as prostitution? Do we violate fundamental rights by permitting the practice?[5] Do we violate fundamental rights by forbidding it?[6] What about the sale of manual labor that is injurious to the body?[7] In other words, there is deep disagreement about whether the right over one's body and its uses is a property-like right that must be alienable in exchange for money, or is a dignity or privacy-like right that must not be sold.

In this article, I argue that both sides are off the mark. The concept of a monolithic, fundamental right to bodily "integrity" is both descriptively and normatively wrong. There should be no legal "right to control one's own body," saleable or not, with a scope that matches up perfectly with the physical borders of the organic, physically continuous human body. This is not to say that we should abandon many familiar rights such as the right to not be tortured or raped. Rather, it is to say that there should be no one-to-one mapping between the physical borders of the organic, integrated human body and the legal borders of the rights derived from it.

For instance, once we determine the relationship between our bodies and fundamental rights, we might not derive any freedom to resist vaccination (although certainly other rights, such as religious rights, might protect this freedom[8]). On the other hand, we might be led to protect freedoms of dress and makeup, even though these activities do not

---

5. *See* Stephanie Farrior, *The International Law on Trafficking in Women and Children for Prostitution: Making it Live Up to Its Potential*, 10 HARV. HUM. RTS. J. 213, 218 n.24 (1997) (noting the argument made by Coalition Against Trafficking in Women that prostitution "usurps and negates prostitute women's right to human dignity, bodily integrity and physical and mental well-being"); Jane E. Larson, *Prostitution, Labor, and Human Rights*, 37 U.C. DAVIS L. REV. 673, 677 (2004) (pointing out that the view—"any form of prostitution is a human rights violation"—is supported by the 1949 UN Convention for the Suppression of the Traffic in Persons and the Exploitation of the Prostitution of Others where it states that "prostitution and . . . traffic in persons for the purposes of prostitution are incompatible with the dignity and worth of the human person").

6. *See* Int'l Comm. for Prostitutes' Rights, *World Charter for Prostitute Rights*, *in* A VINDICATION OF THE RIGHTS OF WHORES 40, 40 (Gail Pheterson ed., 1989), *available at* http://www.walnet.org/csis/groups/icpr_charter.html; Norma Jean Almodovar, *For Their Own Good: The Results of the Prostitution Laws as Enforced by Cops, Politicians and Judges*, 10 HASTINGS WOMEN'S L.J. 119, 123 (1999) (arguing that the subsidiary effects of criminalizing prostitution is "violative of human rights and dignity"); Larson, *supra* note 5, at 681 (describing arguments for legitimizing prostitution on the grounds that it "is a free choice made by an autonomous individual" and that "[r]espect for women's self-determination requires respect for female choices about sex and survival"); *see also* DEBORAH L. RHODE, JUSTICE AND GENDER: SEX DISCRIMINATION AND THE LAW 262 (1989).

7. *See, e.g.*, RABBINICAL COUNCIL OF AM., POLICY STATEMENT ON HUMAN DIGNITY AND LABOR (1987), http://www.rabbis.org/news/article.cfm?id=101063 (describing the fact that "workers in chemical factories may be exposed to dangerous chemicals or radiation, the effects of which are not recognized or diagnosed for many years" as compromising human dignity, to explain a position supporting "legislation that will raise the standards of employment" in that and other industries).

8. *E.g.*, Sherr v. Northport–East Northport Union Free Sch. Dist., 672 F. Supp. 81, 91 (E.D.N.Y. 1987) (finding that under the First Amendment, religious exemption to New York's mandatory inoculation program for school children must be extended to all persons who sincerely hold religious beliefs, and not just persons who are bona fide members of a recognized religious organization).

involve the manipulation of one's organic, physically continuous human body.[9] Similarly, the relationship of the body to fundamental rights might lead us to regulate contracts entered into between patients and sellers of prostheses, despite the fact that prostheses are not human. On the other hand, we might not find a justification for either prohibiting or permitting blood sales.[10]

In Part I, I argue that those who promote an inalienable or noncommodifiable right to bodily integrity—a right that cannot be exchanged for money—fail to recognize the fact that possession is not the only important right in a physical object. Even when the sale of possessory rights in body parts is prohibited, the sale of use and exclusion rights in the body, such as the sale of manual labor, are permitted, and these sales can cause the same sort of physical harm and risk that sale of possessory rights can cause. Indeed, I show that even market exchanges in the possession and use of *non-human* objects can cause the same sort of physical harm and risk that property rights in parts and uses of the human body can cause. This is increasingly the case given dependence on medical and other technological devices for sustaining life and health. In sum, those promoting dignity-based rights in the organic, integrated, human body that cannot be alienated for money are descriptively wrong about the body's complex relationship to its surroundings.

In Part II, I argue that those who favor a property-like right to possess and use the human body and its parts free from limits neglect the fact that property and contract rights can be regulated: ownership of an object does not entail sole dominion over the object. Thus, even if "your body is your property," that does not entail a fundamental right to free markets in bodies, their parts, and their uses. In other words, those promoting property-like rights in the organic, integrated, human body that must be alienable for money, as a matter of fundamental rights, are descriptively wrong about property.

In Part III, I argue that as a normative matter, we should not use dignity or autonomy arguments to promote any monolithic right to bodily integrity, whether alienable or not. However, my critique of a monolithic concept of a right to bodily integrity, demarcated by the organic, human, physically integrated body's borders, still leaves room to justify a more nuanced version of rights in the body. Thus, I consider dignity and autonomy as approaches to justifying more nuanced rights in the human

---

9. *Cf.* Gowri Ramachandran, *Freedom of Dress: State and Private Regulation of Clothing, Hairstyle, Jewelry, Makeup, Tattoos, and Piercing*, 66 MD. L. REV. 11, 60–69 (2006).

10. However, there may be other, purely policy-based, non-rights-based justifications for the regulation of those transactions, such as public health. *See, e.g.*, RICHARD TITMUSS, THE GIFT RELATIONSHIP (1970), *reprinted in* THE GIFT RELATIONSHIP: FROM HUMAN BLOOD TO SOCIAL POLICY, 57–315, at 314 (Ann Oakley & John Ashton eds., 1997) ("[C]ommercial markets are much more likely to distribute contaminated blood; the risks for the patient of disease and death are substantially greater.").

body, ones whose legal borders may not line up with the human body's physical borders. I conclude that dignity as a basis for singling out the body and prohibiting these sales is too arbitrary for a pluralistic society. Autonomy as a basis for prohibiting the regulation of these sales ignores the inevitable presence of economic and cultural coercion. Finally, I consider and critique the capability approach to justifying nuanced rights in the human body, since it arguably solves many of the problems identified with dignity and autonomy.

Nevertheless, all this critique need not bring us to the conclusion that human or fundamental rights have nothing to say about torture, prostitution, or forced labor. In Part IV, I propose that the body's relationship to fundamental rights can be rooted in a proceduralist approach—such as the use of fundamental rights to promote a dynamic, evolving culture—just as speech rights are thought to promote political and cultural processes. The reason is that the body is uniquely "personal and political."[11] The body is especially personal because it grounds our subjective experience, the experience that others cannot directly access. But the body is also especially political because it is a primary site for exploring different values, subcultures, and identities. This means that to avoid a stagnant culture, we should use fundamental rights to avoid the consolidation of control over individuals' bodies, whether in the hands of the state or in the hands of economically powerful actors. Avoiding such consolidated control is a key component of ensuring that the formation of identities, subcultures, and cultural values is not directed by a powerful few.

This way of justifying the relationship between the body and fundamental rights does not lead to either an inalienable "right to bodily integrity" or to a particularly alienable "right to control one's own body." Under this approach, the organic, integrated body no longer marks any border between what we have human or fundamental rights in and what we do not. To demonstrate, in Part VI provide some tentative examples of how we would apply this view to questions such as state intrusion on the body, rights to modify one's own body, prostitution, organ selling, and the regulation of property and contract due to its effect on bodies, such as patent law, employment and housing law, and welfare law. Instead of creating fundamental rights to protect "bodily integrity" in these situations, we should create fundamental rights where necessary to avoid monopolies over bodies.

---

11. "The personal is political" was coined by members of New York Radical Women, including Carol Hanisch, in the late 1960s. JENNIFER BAUMGARDNER & AMY RICHARDS, MANIFESTA: YOUNG WOMEN, FEMINISM, AND THE FUTURE 19 (2000).

I. WHAT THE DIGNITY OR PRIVACY VERSION OF BODILY INTEGRITY IS MISSING DESCRIPTIVELY: BODIES DEPEND ON ENVIRONMENTS

There may be simple utilitarian policy reasons to regulate or deregulate particular instances of body commodification. For instance, some argue that prohibiting blood selling improves the quality of the supply of blood.[12] Others argue that permitting prostitution would increase economic and social welfare.[13] But in this article, I focus on rights in bodily use and possession that would trump these typical political concerns of increased economic welfare, public health, or the expression of a society's cultural values through law.[14] In other words, should constitutional law, human rights law, or other forms of counter-majoritarian law have any role in these debates? If so, what should those rights in bodily use and possession look like?

There is a great deal of disagreement over this question, but one significant class of views on the subject relies on the following concern: the commodification of the human body's uses or components risks harm to human dignity, or perhaps privacy or personhood, in a way analogous to the way that slavery, the ultimate commodification of the human body, does.[15]

For this set of views, affording the formal legal status of property to the body or its uses and parts is problematic because it threatens what is often described as a fundamental right to human dignity.[16] Alternately, it is described as threatening a fundamental right to bodily integrity,[17] respect for the sacredness or sanctity of human life,[18] religious moral man-

---

12. *See, e.g.*, TITMUSS, *supra* note 10, at 206, 270, 314.

13. *E.g.*, Almodovar, *supra* note 6, at 127 ("'Quality of life' is a subjective concept and the least impressive argument for the continued harassment, arrest and incarceration of a group of people who are trying to improve their quality of life by earning a living."); Martha C. Nussbaum, *"Whether from Reason or Prejudice": Taking Money for Bodily Services*, 27 J. LEGAL STUD. 693, 696 (1998) ("The legalization of prostitution, far from promoting the demise of love, is likely to make things a little better for women who have too few options to begin with.").

14. *Cf.* Dan M. Kahan, *The Cognitively Illiberal State*, 60 STAN. L. REV. 115, 115 (2007) (arguing that "individuals naturally impute socially harmful consequences to behavior that defies their moral norms," thus supporting law that "repress[es] morally deviant behavior" while "honestly perceiv[ing] themselves to be motivated only by . . . harm prevention").

15. *See* MICHELE GOODWIN, BLACK MARKETS: THE SUPPLY AND DEMAND OF BODY PARTS 194–96 (2006) (describing authors opposing commodification of body parts on the grounds that it is like slavery); *see also* DAVID B. RESNIK, OWNING THE GENOME: A MORAL ANALYSIS OF DNA PATENTING 1–7 (2004) (describing controversy over DNA patenting that includes slippery slope arguments ending in slavery).

16. *E.g.*, sources cited *supra* note 3.

17. *E.g.*, Rao, *supra* note 2, at 387–88 (describing the American constitutional right to bodily integrity and the fact that it has the character of a privacy interest, not a property interest, in that it protects against state interference but does not protect a right to alienate the interest).

18. *E.g.*, Proclamation No. 8339, 74 Fed. Reg. 3955 (Jan. 15, 2009) (proclaiming the President's designation of January 18, 2009 as National Sanctity of Human Life Day); Daniel C. Dennett, *How to Protect Human Dignity from Science*, *in* HUMAN DIGNITY AND BIOETHICS, *supra* note 3, at 39, 40 ("Human life, tradition says, is infinitely valuable, and even sacred . . . ."); William P. Clark, Op-Ed, *For Reagan, All Life Was Sacred*, N.Y. TIMES, June 11, 2004, at A27, *available at* 2004 WLNR 5467725.

dates,[19] or "personhood."[20] This is part of the justification provided for prohibitions on organ selling[21] and prostitution.[22] It is not just moral conservatives who fit within this camp, however. There are disability rights activists, those who are concerned with the rights of sexual minorities, and egalitarians who oppose, on dignitarian-type grounds, leaving the modification of bodies up to markets.

The argument might go as follows: A ban on slavery means that humans are fundamentally ineligible for the legal status of property—that which persons (including corporations) may both exercise control over and alienate, either in whole or in part, and often in exchange for money.[23] Therefore, to allow for property rights in parts or uses of humans would blur the line between humans and property. This would threaten dignitary rights in our own bodies, respect for our own bodies, and perhaps even respect for human dignity itself.[24] Of course, there are more nuanced objections to commodification of the body,[25] but many of them share a reliance on the concept that some set of commodification practices threatens human dignity, a fundamental aspect of "personhood" such as bodily integrity, or some similar universal concept such as a fun-

---

19. *E.g.*, *Vatican Bishop Points to Modern Social Sins*, CATHOLIC NEWS AGENCY, Mar. 11, 2008, http://www.catholicnewsagency.com/new.php?n=12031 (reporting the inclusion of "'bioethical' violations such as birth control" and "'morally dubious' experiments such as stem cell research" in a list of seven modern social sins released by the Vatican via an interview with Bishop Girotti); *see also* Nicola Gori, *The New Forms of Social Sin*, L'OSSERVATORE ROMANO, Mar. 9, 2008, *available at* http://blog.acton.org/uploads/penitentiary_interview.pdf (interview with Bishop Girotti).

20. *E.g.*, Margaret Jane Radin, *Market-Inalienability*, 100 HARV. L. REV. 1849, 1852, 1885 (1987) [hereinafter Radin, *Market-Inalienability*]; *see also* MARGARET JANE RADIN, CONTESTED COMMODITIES 55–56 (1996) [hereinafter RADIN, CONTESTED COMMODITIES]; Margaret Jane Radin, *Property and Personhood*, 34 STAN. L. REV. 957, 1014–15 (1982) [hereinafter Radin, *Property and Personhood*].

21. *E.g.*, U.S. DEP'T OF HEALTH & HUMAN SERVS., *supra* note 3, at 96; Munzer, *supra* note 2, at 266 (describing the Kantian human dignity argument as applied to organ sales).

22. *E.g.*, Susan Estrich, *What I Could't Teach Spitzer at Harvard Law*, REAL CLEAR POLITICS, Mar. 13, 2008, http://www.realclearpolitics.com/articles/2008/03/what_i_couldnt_teach_eliot_spi.html:

> I really believe it is none of my business, as a member of the public or the media, if a political or business leader has an affair. I don't sit in judgment of other people's marriages or their private lives. But prostitution isn't just sex. Prostitution objectifies the women who engage in it, dehumanizes sex and sexuality, and turns both into commodities with a price tag.

See also Ann Lucas, *The Currency of Sex: Prostitution, Law, and Commodification*, *in* RETHINKING COMMODIFICATION 248, 248 (Martha M. Ertman & Joan C. Williams eds., 2005) ("[M]ost objections to prostitution are commodification-based.").

23. STEPHEN R. MUNZER, A THEORY OF PROPERTY 48–49 (1990) ("Property rights are body rights that protect the choice to transfer."); Radin, *Market-Inalienability*, *supra* note 20, at 1854 n.19 ("Traditional property rights are alienable in all senses except cancellation; they may be forfeited, relinquished, waived, condemned, and transferred by both gift and sale.").

24. *See, e.g.*, DAVID B. RESNIK, OWNING THE GENOME: A MORAL ANALYSIS OF DNA PATENTING 1–7 (2004) (describing the controversy over DNA patenting that includes "slippery slope" arguments ending in slavery).

25. *See, e.g.*, Rao, *supra* note 2, at 455–56 (arguing that a principle in accord with our intuitions on the subject would allow for property rights in body parts that have been detached from a person, but would not allow for property rights in parts still internal to a person's body).

damental human "capability"[26] for bodily integrity. The idea has roots in Catholic thinking,[27] and also appears in a Kantian argument that human dignity requires persons not to make these sales.[28] Bioethics committees and advisory boards commonly use the fact that something involves body commodification as a reason to be ethically concerned about it.[29]

The public approach to this question—whether commodification of the body is *per se* problematic—is somewhat schizophrenic, especially in the United States. Selling of organs is forbidden by statute,[30] and blood banks will not keep sold blood; yet some plasma centers use purchased blood.[31] The sale of manual labor is permitted, including physically intimate manual labor such as massage therapy. Even payment to volunteers who are willing to have their colons and vaginas invaded by medical students learning how to perform various exams is permitted.[32] However, prostitution is forbidden in most of the United States.

Nevertheless, arguments against various practices of body commodification, *on the grounds* that commodification of the body is problematic to dignity or personhood, continue to have a great deal of purchase, despite the existence of practices ranging from plasma selling to the largely uncontroversial sale of physical labor.[33] The view that there is something threatening about assigning the legal status of property to people or parts of people, at least some of the time—for instance, when the part of the person is not detached[34]—continues to have a great deal of

---

26. NUSSBAUM, *supra* note 1, at 78 (listing bodily integrity as one of ten "central human functional capabilities"). It is important to note that although Nussbaum supports an intuitionist, seemingly dignitarian approach to developing her list of capabilities, and although bodily integrity maintains a place on her list, her specific views on how to apply those capabilities to law do not necessarily lead to anti-body-commodification positions. For instance, Nussbaum has elsewhere argued for the legalization of prostitution. *See* Nussbaum, *supra* note 13, at 723–24. On the other hand, Radin has derived from Nussbaum's capabilities approach and concept of human flourishing a stance against some forms of commodification. RADIN, CONTESTED COMMODITIES, *supra* note 20, at 64, 93–96.

27. *See* U.S. CONFERENCE OF CATHOLIC BISHOPS, SHARING CATHOLIC SOCIAL TEACHING 5 (1998), *available at* http://www.archdiocese-chgo.org/departments/peace_and_justice/pdf/teaching_doc/sharing_social_t.pdf ("[T]he Catholic Church proclaims that human life is sacred and that the dignity of the human person is the foundation of a moral vision for society. . . . [This belief] is the foundation of all the principles of our social teaching.").

28. *See* Munzer, *supra* note 2, at 259; Shell, *supra* note 3, at 344–45.

29. *See, e.g.*, U.S. DEP'T OF HEALTH & HUMAN SERVS., *supra* note 3, at 96; *see also* The President's Council on Bioethics, http://www.bioethics.gov (last visited Nov. 2, 2009) (listing "Property in the Body" as one of its main "Topics of Council Concern"). The Council was disbanded on June 11, 2009. Nicholas Wade, *Obama Plans to Replace Bioethics Panel*, N.Y. TIMES, June 18, 2009, at A24, *available at* 2009 WLNR 11613820.

30. 42 U.S.C. § 274e(a) (2006).

31. Steven R. Salbu, *AIDS and the Blood Supply: An Analysis of Law, Regulation, and Public Policy*, 74 WASH U. L.Q. 913, 942 n.163 (1996) (describing the move in the United States towards an all donated blood supply, as encouraged by FDA regulations).

32. *See* Nussbaum, *supra* note 13, at 701, 706.

33. *See id.* at 693 ("ALL of us, with the exception of the independently wealthy and the unemployed, take money for the use of our body."); Salbu, *supra* note 31, at 944–46.

34. *See, e.g.*, Rao, *supra* note 2, at 455–56 (arguing that a rule forbidding property rights in body parts when the parts are internal to a person, but permitting property rights in body parts when detached from a person, would fit best with our intuitions).

force. Even anxiety about prostitution, which involves the sale not of a body part, but rather of a service employing the physical body, is often expressed in the terms that it is a commodification of the body.[35]

However, it is not the specter of human parts or even humans having the formal legal status of property that makes slavery and indentured servitude so horrifying, or markets in human body parts so worrisome.[36] In other words, we should not rest our concern with these practices on a formalist idea that humans or human parts having the legal status of property is itself an affront to some definition of what it means to be a human, to some concept of inviolable dignity. This focus is descriptively misplaced, as it relies on an overly simplistic conception of the relationship between humans' bodies and their environments. Whatever human dignity or personhood means, property rights *already* threaten it, even where the line between property and persons is clear.

For instance, if we prohibited the sale of possession rights in bodies and their parts, the sale of use rights would still pose risks to dignity in the sense of health, physical well-being, and physical independence: "The precedents for making use of one another's bodies are sufficiently well established in manual labor . . . ."[37] But manual labor often causes changes to the body over time, such as repetitive stress injuries. Suppose person A sells use of her hands until their function is so impaired that she requires assistance to get dressed and lives with constant pain. Suppose person B sells possession of her blood, or even a kidney, but remains healthy and relatively pain-free. Although person A "merely" sold the use of her body, and person B sold a possessory right to part of her body, can we say that the laborer has somehow risked or harmed her bodily dignity less?

On the other hand, it would be infeasible to go so far as to prohibit sales of use rights in the body because this would prohibit all wage labor. The use of person A's body was commodified not just when she was injured, but even when she sold the "labor of animation."[38] And a professor commodifies the use of her body by agreeing to stand in front of a

35. *E.g.*, sources cited *supra* note 22.

36. *See* GOODWIN, *supra* note 15, at 198:

> Slavery's pernicious effect is not exclusively derived from a market evaluation in the human body. Indeed, many slaves were given away as gifts. . . . The villainy of slavery is best characterized by the creation of a chattel system wherein Black men, women, and children were explicitly and exclusively exploited; stripped of their humanity, tortured, bred, denied legal protection, forbidden educational instruction and religious expression.

37. Martyn Evans, *The Utility of the Body*, *in* OWNERSHIP OF THE HUMAN BODY: PHILOSOPHICAL CONSIDERATIONS ON THE USE OF THE HUMAN BODY AND ITS PARTS IN HEALTHCARE 207, 207 (Henk A.M.J. Ten Have & Jos V.M. Welie eds., 1998).

38. Sarah S. Jain, *The Prosthetic Imagination: Enabling and Disabling the Prosthesis Trope*, 24 SCI., TECH., & HUM. VALUES 31, 32 (1999) (internal quotation marks omitted) (quoting Elaine Scarry, *The Merging of Bodies and Artifacts in the Social Contract*, *in* CULTURE ON THE BRINK 85, 97 (Gretchen Bender & Timothy Druckery eds., 1994)).

classroom for specified periods of time and talk about particular subjects, using her mouth, lungs, and brain.[39]

One might imagine that we can separate the professor example from the prostitute or someone engaging in manual labor based on whether the work is typically thought of as selling the "use of one's body"—whether the work is culturally viewed as bodily work. But even work not typically thought of in this way could raise concerns about bodily dignity. For example, an employer who requires an employee to wear particular clothes has not in fact required the employee to change his organic, physically continuous body. Most would not describe such an employee, presuming he is not a model, as having sold the use of his body. However, the clothes worn may have an effect on the employee's body so significant that it cannot be cognitively overcome. Suppose the clothing required is high heels. Walking about in high heels is an entirely different bodily experience than not, even if we leave aside the possibility of long term injury to the body that can stem from wearing high heels. Or suppose the clothing required is tight jeans, skirts with petticoats, or other physically uncomfortable or cumbersome attire; is this somehow less physically intrusive or cognitively overpowering to the worker than the process of having one's blood withdrawn for sale, or an employer's requirement that one's fingernails be trimmed? Umberto Eco has described the experience of wearing tight jeans as so physically affecting that it makes him write differently.[40] For many persons, the tight jeans requirement is far more intrusive than the fingernail-trimming requirement, and for some, it would even be experienced as more intrusive than a requirement to obtain a vaccination.

It is not just sales of the use of one's body that can cause the same sorts of physical harm or dignitary intrusions that sale of the parts of one's body can. Even the exchange of use or possession rights in completely inorganic material which has no attributes of personhood can result in the same type of threats to bodily integrity or self-control that property rights in human parts are thought to raise. For instance, it is completely unremarkable for property rights to exist in electronic gadgets. But we might be concerned if owners of patents on products such as pacemakers and robotic arms were permitted to enforce "end user license agreements" ("EULAs")[41] against patients. These EULAs could in effect restrict what patients can do with products that have become merged with their own bodies. And we should rightly, I argue, be similarly con-

39. Nussbaum, *supra* note 13, at 704 (constructing this example).

40. UMBERTO ECO, *Lumbar Thought*, *in* TRAVELS IN HYPERREALITY 191, 193–95 (William Weaver trans., Harcourt Brace Jovanovich 1986) (1983).

41. An End User License Agreement is a contract that dictates the terms under which an end-user can use software or a device. Users often become a party to these contracts without reading them due to their length and form. *See* Kevin W. Grierson, Annotation, *Enforceability of "Clickwrap" or "Shrinkwrap" Agreements Common in Computer Software, Hardware, and Internet Transactions*, 106 A.L.R. 5TH 309, § 2(a) (2003).

cerned with the effects of property rights in wheelchairs, cochlear implants, tools used in labor, and other such devices on the bodies of those who need or desire to use them.

If that example is too sui generis, we can consider that exercise of even the most mundane property right—the right to develop real estate—can affect the bodily dignity of others. If that property right is unregulated with respect to the installation of wheelchair ramps, then how property holders choose to exercise the right will have far-reaching effects on the mobility and lived bodily experience of persons who require wheelchairs to get around. Whether such persons' bodies are "disabled" or not will depend on how those property rights are exercised. And yet, the property holders will have had this effect on others' bodies without having physically "touched," possessed, or used the bodies of the persons using wheelchairs.[42]

In other words, we can construct many examples where someone has exercised property-like rights over something other than possession of a person or its parts, yet has entailed significant effects on another's lived physical experience. Some of the examples involved the exercise of possession and use rights over non-human objects, while other examples involved the sale of use rights in a human or its parts. In some of these examples, the effects on lived physical experiences are even more significant than the effects of selling possessory rights to a part of the body—selling blood, for instance. The reason we can construct these examples is that bodies, like any other objects, have a complex relationship with their environment.[43] Just as failure to provide a neighbor with an easement can affect the neighbor's ability to use his property,[44] failure to install a wheelchair ramp can affect another's ability to use his body.

This is why refraining from physical intrusions on the human body does not necessarily match up with maximal dignity for the person. There are many ways for the dignity or integrity of bodies to be very significantly affected by cultural norms, markets in items other than bodies, and political rules. Once we recognize that our physical and social

---

42. *See* Samuel R. Bagenstos, *Subordination, Stigma, and "Disability,"* 86 VA. L. REV. 397, 429 (2000) (describing an acceptance of the social model of disability in the United States, in which "disability is attributed primarily to a disabling environment instead of bodily defects or deficiencies" (internal quotation marks omitted) (quoting Harlan Hahn, *Feminist Perspectives, Disability, Sexuality and Law: New Issues and Agendas*, 4 S. CAL. REV. L. & WOMEN'S STUD. 97, 101 (1994)); Mairian Corker & Tom Shakespeare, *Mapping the Terrain*, *in* DISABILITY/POSTMODERNITY: EMBODYING DISABILITY THEORY 1, 2–3 (Mairian Corker & Tom Shakespeare eds., 2002) (describing the shift from medical model of disability to social model, in which activists raised awareness of the fact that impairment alone does not cause disability, but rather social and economic conditions overlaying impairment).

43. *See* sources cited *supra* note 42.

44. For instance, an easement by necessity can be required to ensure that another's parcel of land has road access. *E.g.*, Wilson v. Smith, 197 S.E.2d 23, 25 (N.C. Ct. App. 1973) ("A way of necessity arises when one grants a parcel of land surrounded by his other land, or when the grantee has no access to it except over the land retained by the grantor or land owned by a stranger.").

experience of our bodies cannot be easily isolated from the environment, we can see why simply drawing a physical line between persons and property should not satisfy any concerns we might have with bodily dignity, even if we could agree on what dignity means. We should not be satisfied that we have protected persons physically by defining personhood so as to include, as most accounts do, bodily integrity.[45] As a descriptive matter, this approach is too simplistic.

While the fiction of a human body that could ever be made totally free from "intrusion" or given complete "integrity"[46] continues to retain purchase in debates about activities such as organ selling and prostitution,[47] it is about to face a profound challenge: developments in biotechnology have already begun to result in very obvious physical mergers of human and non-human parts and entities. For instance, in the case of a pacemaker or a robotic arm, humans are merging with inorganic property of the sort that is routinely and uncontroversially commodified.[48] In the case of merging plant genetic material with human genetic material, humans may merge with property that is organic, but which is nevertheless still routinely commodified.[49] In the case of merging human and animal genetic material, transplanting an animal organ into a human, or growing a human organ inside an animal, humans may merge with organic property that is legally, but controversially, commodified.[50] In some ways this merger is nothing new.[51] As I've argued above, very common forms of property such as real estate can have significant effects on the lived physical experience of our own bodies.[52] However, developments in biotechnology are making this fact more visible. In the long run, as these mer-

---

45. *E.g.*, JAMES HUGHES, CITIZEN CYBORG: WHY DEMOCRATIC SOCIETIES MUST RESPOND TO THE REDESIGNED HUMAN OF THE FUTURE 73, 229–31 (2004) (arguing for "personhood-based 'cyborg citizenship' versus 'human racism,'" yet nevertheless asserting that personhood requires self-ownership of bodies); NUSSBAUM *supra* note 1, at 78 (including "bodily integrity" in her list of capabilities).

46. *See* DONNA J. HARAWAY, *A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century*, *in* SIMIANS, CYBORGS, AND WOMEN: THE REINVENTION OF NATURE 149, 150 (1991); Jain, *supra* note 38, at 43.

47. *See, e.g.*, sources cited *supra* notes 3, 21, 22.

48. *E.g.*, Brian Mockenhaupt, *Rebuilding: Bryan Anderson*, ESQUIRE MAG., Mar. 2008, at 184, *available at* 2008 WLNR 25444912.

49. *See generally* Michael D. Rivard, *Toward a General Theory of Constitutional Personhood: A Theory of Constitutional Personhood for Transgenic Humanoid Species*, 39 UCLA L. REV. 1425, 1434–41 (1992) (describing processes for creating transgenic species).

50. *See* Elizabeth L. DeCoux, *Pretenders to the Throne: A First Amendment Analysis of the Property Status of Animals*, 18 FORDHAM ENVTL. L. REV. 185, 220–21 (2007) (describing the successful implantation of human embryonic stem cells into mouse embryos, and the resulting mouse–human hybrids, which had a small amount of human brain cells fully integrated into their brains otherwise made up of native mouse brain cells).

51. Some argue that humans are "embodied in an extended technological world," rather than existing as "distinct beings in an antagonistic relationship with their surroundings." Robert Pepperell, *The Posthuman Manifesto*, KRITIKOS, Feb. 2009, http://intertheory.org/pepperell.htm; *cf.* HARAWAY, *supra* note 46, at 150 (attempting to disrupt the naturalization of the human body and of gender, but without seeking to be innocent or transcendent of the systems of power and coercion that socially construct gender and the body).

52. *See supra* text accompanying note 42.

gers between people and property become more common, sustaining different formal legal statuses for people and other objects may appear more and more arbitrary.

Prostheses such as robotic arms are likely to become common in part because the wars in Iraq and Afghanistan have produced hundreds of amputees, spurring large increases in funding for prosthetic technology. Amazing developments in the field are occurring.[53] Researchers at the University of Pittsburgh, for instance, have recently announced the successful development of a robotic arm that a monkey can learn to control sufficiently to pick up and eat food, using only a small device installed in its brain.[54]

Transgenic species, or entities which include DNA from two different species, have already been created.[55] The species appear to exhibit phenotypic characteristics of both species.[56] Animal-human "hybrids" or plant-human "hybrids"[57] might someday be useful for growing or developing replacement organs or tissue for humans, as one example of the drive behind this research.[58]

This prominent merger of humans and property has raised a great deal of anxiety over how we may avoid the state of legalized slavery or other related threats to human dignity. For instance, recent scholarship on "patenting people" has concerned itself with the mixture of human and animal parts, especially genetic material.[59] This scholarship expresses fear that because human genetic material can currently be patented legally, and has been, eventually animal-human hybrids with many human characteristics may be patented, and perhaps even turned into a slave class.[60] The specter now exists of a farm of highly intelligent pig-human

---

53. *See* Mockenhaupt, *supra* note 48, at 184 ("Funding for prosthetics projects has swelled significantly over the past six years. The Defense Advanced Research Projects Agency, or DARPA, is bankrolling one of the most ambitious programs, the $85 million Revolutionizing Prosthetics project to build a prosthetic arm that matches the human arm in functionality by next year. . . . The fusion of man and machine is upon us, the extraordinary enabling the everyday.").

54. Benedict Carey, *Monkey Thinks, Moving Artificial Arms as Own*, N.Y. TIMES, May 29, 2008, at A1, *available at* 2008 WLNR 10097697.

55. *See* DeCoux, *supra* note 50, at 222.

56. *Id.*

57. Politicians have referred to animal–human hybrids, probably meaning both chimeras and transgenic species, where chimeras contain cells from two different entities, and transgenic beings contain DNA from two different entities. *See* Stephen Munzer, *Human–Nonhuman Chimeras in Embryonic Stem Cell Research*, 21 HARV. J.L. & TECH. 123, 124 (2007) (providing these definitions).

58. *See* Bagley, *supra* note 2, at 505–06.

59. *Id.* at 506–07.

60. *See, e.g., id.* at 502, 511–12. Indeed, in 2006, Cardozo Law School hosted an entire two-day conference on the legal, moral, and policy implications of patenting human DNA, human–animal hybrids, and other biotechnological innovations. Numerous law professors, government patent lawyers, and others presented arguments at this conference, at which Bagley was the keynote speaker. *See* Margo A. Bagley, Keynote Address at the Benjamin N. Cardozo School of Law Patenting People Conference (Nov. 12, 2006).

hybrids being raised for removal of their kidneys and slaughter.[61] Property rights in human genetic material are thought to also threaten humans' rights to do as they will with their own bodies—their own genetic material. A patent-holder on genetic material might, for instance, prevent a person with that material from reproducing.[62]

Some of the literature has drawn on these fears to simply oppose the technology that blurs the line between humans and property. The opposition ranges from proposing outright bans to moderate "think before we act" proposals.[63] This opposition literature finds itself in alliance with religious conservative activists who also oppose technologies that blur the line between humans and non-humans.[64] These religious activists argue that without a rigid definition of humans, encompassing everything from embryos to persons in irreversible vegetative states, the degradation of human dignity will eventually result.[65]

In contrast, some argue that we must simply update our definition of what a legal person is to potentially include entities that are not completely human.[66] On this view, armed with a proper definition of legal personhood, we could simply forbid property rights in anything that counts as a person, whether completely human or not: property rights in an embryo would be fine perhaps, but property rights in a pig with a high functioning human brain would perhaps not be acceptable.

But despite the existence of these dramatically different responses, both fail to come to terms with the hard truth that all of this technology confronts us with: it is not enough to decide the difficult question of who is a person and then simply forbid property rights in that person. It takes much more than giving persons' bodies a special, formal legal status of "non-property" to sufficiently protect the physical well-being and independence of those bodies.

II. WHAT THE PROPERTY VERSION OF BODILY INTEGRITY IS MISSING DESCRIPTIVELY: PROPERTY IS NOT SOLE DOMINION

An alternate approach to the question of what bodily integrity might mean views control over one's body as a property-like right, essential to freedom or autonomy. In this camp are persons who believe that there is

---

61. *Cf.* Bagley, *supra* note 2, at 506 (describing a patented method for producing human organs from pigs).

62. *See* HUGHES, *supra* note 45, at 231 (proposing regulation to prevent this from happening).

63. *See* Bagley, *supra* note 2, at 473.

64. *E.g.*, *id.* at 510 (quoting with approval Professor Leon Kass, a religiously-inspired bioethicist and former chair of President George W. Bush's Council on Bioethics).

65. *See id.* at 496–97.

66. As an example of this approach, the U.S. Patent and Trademark Office initially rejected a patent application for a HuMouse because it "embraces a human being," arguing that the Thirteenth Amendment would prohibit such a patent. Gregory R. Hagen & Sébastien A. Gittens, *Patenting Part-Human Chimeras, Transgenics and Stem Cells for Transplantation in the United States, Canada, and Europe*, 14 RICH. J.L. & TECH. 11, 34, 49 (2008).

a fundamental right to control over one's own body or its parts.[67] This theory would not only protect the body from unwanted intrusion, but also would protect one's right to modify one's body, choose to accept or reject medical treatment, and the like. However, autonomy includes the autonomy to contract one's autonomy away. Thus, the strong version of this view would require that commodification of the body be permitted, not simply when it is good economic or social policy, but because to do otherwise would interfere with fundamental rights to ownership of one's own body.[68] Strong libertarians would see prohibitions on the commodification of human bodies as unjust impositions on autonomy, and a coherent application of this view would therefore seem to permit organ selling, prostitution, the sale of genetically prized embryos, or employer preferences for "upgraded" workers.[69]

But to the extent this view derives the right to commodify one's body from a property right in one's own body,[70] it relies on a descriptively impoverished view of property as "sole dominion" over an object. The person with ownership of the property, on this view, has the right to use it, exclude others, possess it, and alienate it.[71] Because persons own their own bodies, the implication is that they must have the right to sell use and possessory rights in their body and its parts.

---

67. Though they are not libertarians, major proponents of the transhumanist movement, which supports the use of technology to make humans "better than well," have cited this argument from autonomy in support of permitting such technologies. HUGHES, *supra* note 45, at 207, 229 (arguing that "[o]nly beings with personhood are exempt from being property, since we each own ourselves and can't be alienated from ownership of ourselves," and noting, with approval, "disparate movements, like transgender rights, working to radicalize our control over our own bodies"); Nick Bostrom, *In Defense of Posthuman Dignity*, 19 BIOETHICS 202, 210 (2005) ("A liberal democracy should normally permit incursions into morphological and reproductive freedoms only in cases where somebody is abusing these freedoms to harm another person."); *cf.* Volokh, *supra* note 4, at 1835 (arguing for a right to sell and buy organs rooted in the right to self defense). *See also* Bonnie Steinbock, *Sperm as Property*, 6 STAN. L. & POL'Y REV. 57, 66 (1995) (arguing that "individual autonomy should prevail, and sperm is correctly regarded as property that can be bequeathed by will"); Mary Taylor Danforth, Current Topic in Law and Policy, *Cells, Sales, and Royalties: The Patient's Right to a Portion of the Profits*, 6 YALE L. & POL'Y REV. 179, 191–95 (1988) ("Th[e] property-based notion of control over one's name or likeness should be extended to one's body parts. The common law recognizes the individual's exclusive right to use what is inherently personal, and nothing is more personal than one's own genetic material."); Roy Hardiman, Comment, *Toward the Right of Commerciality: Recognizing Property Rights in the Commercial Value of Human Tissue*, 34 UCLA L. REV. 207, 218 (1986) (arguing that "human tissue possesses characteristics that satisfy many of the criteria for establishing rights in tangible property").

68. *See* sources cited *supra* note 67.

69. *E.g.*, DAVID A. J. RICHARDS, SEX, DRUGS, DEATH, AND THE LAW 84–127 (1982) (arguing that properly understood autonomy would lead to permission for prostitution); *cf.* RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 139 (3d ed. 1986) (suggesting that a market in babies could benefit those who can bear children but do not want them and those who cannot bear children but want them).

70. *See* sources cited *supra* note 67.

71. *See* WILLIAM BLACKSTONE, 2 COMMENTARIES *2; *see also* Robert P. Burns, *Blackstone's Theory of the "Absolute" Rights of Property*, 54 U. CIN. L. REV. 67, 69 (1985) (arguing that the reading of Blackstone as promoting "absolute" property rights is overbroad, and that he believed property was "'independent of civil institutions' only in the most theoretical sense").

But descriptively, ownership does not entail sole dominion over property.[72] Instead, property is best thought of as a "bundle of sticks," where the sticks are rights and responsibilities that can be disaggregated from the bundle.[73] Thus, for instance, the law of nuisance may prevent me from using my property in a manner that pollutes the air and disturbs my neighbor.[74] Or, my neighbor may have an easement by necessity across my property, requiring me to permit her to cross my property so that her parcel is not landlocked.[75] She may have this right even as I retain possession of and title to the property. The removal of these "sticks" from my "bundle" of property rights is not generally considered a violation of any fundamental right to my property.[76] That is why even the common law of property includes rights such as the easement by necessity or the nuisance action.[77]

What I do with my property affects other people, and thus, the fact that the body may be analogized to property does not exempt it from this general principle which permits property regulation for the sake of social welfare and efficiency. While extremely low levels of regulation, or even no regulation, might in fact be the best way to achieve social welfare, this social welfare debate over regulation has nothing to do with fundamental rights. For instance, if my body is like any other form of property, whether the state should decriminalize prostitution would depend on whether (1) the practice is inherently unsafe, (2) women are better off overall when the practice is prohibited or when it is permitted, and (3)

---

72. *See* Kristen A. Carpenter, Sonia K. Katyal & Angela R. Riley, *In Defense of Property*, 118 YALE L.J. 1022, 1026–29 (2009) (contrasting the classic ownership model of property with more modern relational models of property, such as a proposed stewardship model, that would be more descriptively and normatively appropriate to recognizing cultural property claims); David Fagundes, *Property Rhetoric and the Public Domain*, 94 MINN. L. REV. (forthcoming 2010) (manuscript at 25, on file with author) (describing the popular myth of property as dominion over things, and contrasting it to the more accurate conception of property as a "system that structures social relationships with resources").

73. *See* Wesley Newcomb Hohfeld, *Fundamental Legal Conceptions as Applied in Judicial Reasoning*, 26 YALE L.J. 710, 742 (1917):

> In opposition to the ideas embodied in the passages just given, it is submitted that instead of there being a single right with a single correlative duty resting on all the persons against whom the right avails, there are many separate and distinct rights, actual and potential, each one of which has a correlative duty resting upon some one person.

74. *See, e.g.*, Aldred's Case, 9 Co. Rep. 57b, 58a (1610) (finding the building of a pig sty too close to a neighbor's house to violate the neighbor's rights).

75. *See, e.g.*, Wilson v. Smith, 197 S.E.2d 23, 25 (N.C. Ct. App. 1973) ("A way of necessity arises when one grants a parcel of land surrounded by his other land, or when the grantee has no access to it except over the land retained by the grantor or land owned by a stranger.").

76. For instance, regulation that limits the use of property is not considered a "taking" in the United States even if discrete segments of the property rights are essentially removed, such as ability to exploit airspace above land. *See* Penn Cent. Transp. Co. v. New York City, 438 U.S. 104, 130 (1978). Nor, for instance, has labor regulation been considered to violate fundamental substantive due process rights to property or contract in the United States since 1937. *See, e.g.*, NLRB v. Jones & Laughlin Steel Corp., 301 U.S. 1, 49 (1937) (upholding the National Labor Relations Act).

77. *See Wilson*, 197 S.E.2d at 26; *Aldred's Case*, 9 Co. Rep. at 58b.

regulation would promote public health and safety better than criminalization.[78]

One might respond to this descriptive point about the legitimacy of property regulation by arguing that the body is a particularly special repository of autonomy. On this view, the body is not like other forms of property, and deserves to be protected by a kind of fundamental right to sole dominion that other property does not entail. A strong proponent of liberalism might go even further and argue that the state should not regulate traditional forms of property either: the state should not prohibit me from being a prostitute *or* from chopping down my trees because both are exercises of my autonomy. But as Part III will show, both of these positions are normatively wrong.

III. WHY DIGNITY, AUTONOMY, AND CAPABILITY APPROACHES TO RIGHTS IN THE BODY ARE NORMATIVELY WRONG

A. What's Normatively Wrong with Dignity

Many may feel that my descriptive critique of the dignitarian approach to bodily integrity is not enough reason to be against a dignitarian right to bodily integrity. Maybe the fiction of a body with integrity that cannot be intruded on is a good fiction—one that should be promoted as much as possible. Perhaps we *should* be carving off the body from the environment in which it exists. For example, some argue that it is a good idea to resist technological advances that will blur the lines between bodies and their environments even further, in order to ensure that people continue to think of human bodies and life as sacred. Thus, perhaps we *should* think of dignity as harmed more severely when the body has the formal legal status of property than when it does not. Perhaps we really should consider the person who sold her kidney to have suffered greater dignitary harm than the person who sold the use of her hands to her employer until they were unusable. Or perhaps we really should consider a prostitute to suffer greater dignitary harm than a person who, as a result of poverty, puts his prosthetic arm up for sale on eBay.

In this section, I argue that attempts to resist technological advances and maintain a rigid distinction between humans and property may lead to less, not more, human and social flourishing. Technology that blurs this line holds the potential to improve the lives of those who are sick, injured, or disabled, or those who simply want to improve and modernize their bodies.

I also critique the use of dignity as a justification for even more nuanced rights in the body, ones that permit some forms of body commodification but not others. The reason is that dignity requires a univer-

78. *See* sources cited *supra* note 13.

sal account of what is, essentially, a human, and what are the components of human dignity.[79] Such universal accounts of what a dignified life consists of are likely biased, and therefore carry with them the danger of becoming oppressive. This is not only a problem with using dignity to promote a monolithic privacy like right to bodily integrity, it is also a problem with using dignity to justify even more nuanced rights in the body.

1. Pluralism

Many have an intuition that a right to control one's own body is fundamental to human dignity, human flourishing, or liberty.[80] This concept might support the right to protect the body from intrusion, but would encompass no right to mutilate one's own body, to consent to euthanasia, sell one's organs, or otherwise show "disrespect" for one's body.[81] The notion that the body has special status is commonly held. For instance, a Department of Health and Human Services task force, in issuing recommendations on organ transplantation, once stated "society's moral values militate against regarding the body as a commodity."[82] It went on to cite a report from the Hastings Center that said:

> The view that the body is intimately tied to our conceptions of personal identity, dignity, and self-worth is reflected in the unique status accorded to the body within our legal tradition as something that cannot and should not be bought or sold. Religious and secular atti-

79. One of the best accounts of this form, and from which we can learn a great deal, is Martha Nussbaum's development and defense of a universal list of basic human capabilities. But her work demonstrates that rigorous attempts to define universal dignitarian human capabilities or rights rely on "intuition," which she acknowledges is subject to "distortion." Our intuitions about what is minimally required for a life with dignity are subject to our own biases. For this reason, Nussbaum is careful to describe the list as contingent and tentative, and also to allow for mechanisms of cultural change that could help us revise the list, such as being able to imagine and reason. MARTHA C. NUSSBAUM, FRONTIERS OF JUSTICE: DISABILITY, NATIONALITY, SPECIES MEMBERSHIP 83, 174–75 (2006).

80. *E.g.*, Vacco v. Quill, 521 U.S. 793, 807 (1997) (clarifying that the right to refuse medical treatment rests in the right "to bodily integrity and freedom from unwanted touching"); Washington v. Glucksberg, 521 U.S. 702, 720 (1997) (citing Rochin v. California, 342 U.S. 165, 172 (1952) (recognizing the right to bodily integrity); Cruzan v. Dir., Mo. Dep't of Health, 497 U.S. 261, 278 (1990) (recognizing "a general liberty interest in refusing medical treatment"); NUSSBAUM, *supra* note 1, at 78 (including it within her list of capabilities); *see also* MODEL PENAL CODE § 211.1 (1980) (consolidating battery, mayhem, and assault into a single crime); RESTATEMENT (SECOND) OF TORTS §§ 13, 18, 21 (1965) (defining battery and assault).

81. *E.g.*, *Vacco*, 521 U.S. at 808 ("By permitting everyone to refuse unwanted medical treatment while prohibiting anyone from assisting a suicide, New York law follows a longstanding and rational distinction.").

82. Henry Hansmann, *The Economics and Ethics of Markets for Human Organs*, 14 J. HEALTH POL., POL'Y & L. 57, 59 (1989) (internal quotation marks omitted) (quoting U.S. DEP'T OF HEALTH & HUMAN SERVS., *supra* note 3, at 96).

> tudes . . . make it plain just how widespread is the ethical stance maintaining that the body ought to have special moral standing.[83]

Many other countries also forbid organ sales.[84] But this view loses its appeal quickly when we consider that what one community may find to express a lack of self-respect, another community may find an ultimately fulfilling and meaningful use of the body.[85] Changing the shape of one's genitals might seem to some "disrespectful" of the body with which one is born, but for some transgender persons, it is a self-affirming expression of self and identity. Having sex in exchange for money may be experienced as degrading for some prostitutes, but can be a fulfilling career for other prostitutes. For instance, some prostitutes defy the presumption that emotional connections cannot be made through commodified sex, by forming friendships with clients, or having "favorite" clients.[86] Which, if any, is most degrading—getting paid to have an advertisement tattooed on one's arm, getting paid to have sex, or getting paid to give a massage? The answer will surely be culturally contingent. Covering one's breasts or hair could be alternately understood as expressing shame or expressing self-respect.[87] In a pluralist society, dignitary concerns seem an inadequate justification for forbidding people from commodifying their bodies—they rely on conceptions of human nature or what constitutes a good life that are not universally or even widely shared, such as religious or "intuitionist" conceptions.

One of the most famous proponents of this dignitarian approach, Leon Kass, the chair of the President's Council on Bioethics from 2002–2005, encourages us to listen to the "wisdom of repugnance" in critiquing certain practices as harmful to human dignity.[88] Kass has lamented not only the immorality of reproductive technologies, but also the offensiveness of licking ice cream cones in public.[89]

---

83. Kevin W. Wildes, *Libertarianism and Ownership of the Human Body*, *in* OWNERSHIP OF THE HUMAN BODY, *supra* note 37, at 143, 152 (quoting U.S. DEP'T OF HEALTH & HUMAN SERVS., *supra* note 3, at 96).

84. *E.g.*, Transplantation of Human Organs Act, No. 42 of 1994 (India); Human Tissue Gift Act, R.S.B.C. 1996, ch. 211 § 10 (Can.). *See generally* WORLD HEALTH ORG., HUMAN ORGAN TRANSPLANTATION: A REPORT ON DEVELOPMENTS UNDER THE AUSPICES OF WHO, 1987–1991, at 15–26 (1991) (surveying and listing countries that ban organ sales and those that do not).

85. *See* Wildes, *supra* note 83, at 147 ("Much of the effort to regulate and constrain the use of the human body is embedded in a moral language of 'sanctity', 'dignity', 'justice', 'integrity', and 'solidarity' that assumes particular moral commitments.").

86. Lucas, *supra* note 22, at 248, 252.

87. *See* FADWA EL GUINDI, VEIL: MODESTY, PRIVACY AND RESISTANCE 13 (1999) (discussing the evolution of differing views pertaining to veiling practices); *cf.* CAROLYN G. HEILBRUN, THE EDUCATION OF A WOMAN: THE LIFE OF GLORIA STEINEM, at xviii (1995) (describing Steinem's "feminis[m] in a miniskirt" as a walking "contradiction").

88. Leon R. Kass, *The Wisdom of Repugnance*, THE NEW REPUBLIC, June 2, 1997, *available at* http://www.catholiceducation.org/articles/medical_ethics/me0006.html.

89. LEON R. KASS, THE HUNGRY SOUL: EATING AND THE PERFECTING OF OUR NATURE 148–149 (1994):

> Worst of all from this point of view are those more uncivilized forms of eating, like licking an ice cream cone—a catlike activity that has been made acceptable in informal

Some members of this camp would permit modifications to the body that do not entail the sale of the body, its parts, or uses, such as the free donation of an organ to a sibling in need, or the creation of an animal-human hybrid over which nobody holds intellectual property rights.[90] Others would object even to some of these modifications that could not fairly be called "commodification"[91] because they involve no economic exchange. But what characterizes this camp is the argument that the root of the right to bodily integrity is a dignitary one. Because people disagree over the contours of dignity, what exactly constitutes an offense against dignity varies widely, and this is why the group contains both moral conservatives and liberals.

This debate over whether a change to or use of one's body is self-respecting appears even when the change is widely thought of as a medical "improvement." For instance, obtaining a cochlear implant that permits the user to hear might seem a thoroughly unobjectionable alteration of one's body to many. But to some members of the deaf community, it can represent a rejection of one's self, as well as one's community—a selling out, if you will, to dominant norms of communication and language.[92] Thus, in a pluralistic society, we cannot rest our rights in the body on what actions demonstrate respect for the body and what actions do not. To do so would rigidly fix the definitions of what are appropriate uses and forms of the body.

Moreover, hewing to convention on the appropriate use and forms of the body will not only lead to stagnation, it lacks compassion. For instance, in his 2006 State of the Union speech, former President George W. Bush opposed animal-human hybrids as one of many technologies that the religious right fears will blur the line between humans and the rest of the world.[93] Even more recently, in March of 2008, the Vatican published a list of seven modern sins that includes genetic manipulation, morally debatable experiments, and violation of the fundamental rights of human nature.[94] But what if animal-human hybrids or stem cell re-

---

> America but that still offends those who know why eating in public is offensive. I fear that I may by this remark lose the sympathy of many readers, people who will condescendingly regard as quaint or even priggish the . . . view that eating in the street is for dogs. . . . This doglike feeding, if one must engage in it, ought to be kept from public view, where, even if *we* feel no shame, others are compelled to witness our shameful behavior.

90. For instance, United States law permits voluntary organ donation while forbidding organ sales. 42 U.S.C.A. § 274e(a) (West 2009).

91. *See id.*; *see also* sources cited *supra* note 19.

92. *E.g.*, SOUND AND FURY (PBS 2000) (documenting a deaf child's wish to obtain a cochlear implant in defiance of her deaf parents' wishes).

93. George W. Bush, Address Before a Joint Session of Congress on the State of the Union (Jan. 31, 2006) (transcript available at http://www.c-span.org/executive/transcript.asp?cat=current&code=bush_admin&year=2006).

94. *See* sources cited *supra* note 19.

search can be used to alleviate a patient's severe pain?[95] Other grounds for opposing these technologies may exist, but a conventionalist's argument from "the wisdom of repugnance"[96] is a particularly uncompelling basis in a pluralistic society. Similarly uncompelling is an argument rooted in the desire to preserve categorical distinctions between persons and property, without a theory of why such a distinction is so worthwhile as to overcome the claims of persons who are physically suffering and their families. Both the kidney seller, who could use compensation to improve her life, and the kidney purchaser, whose health is dramatically improved through transplantation, must be answered to.[97]

2. Dignity of the Body as a Social Justice Concept?

Some argue, in contrast, that compassion actually requires this formalist distinction between people, who are afforded special status, and property.[98] On this view, we must hold the line between persons and property fast, as a way of protecting persons with disabilities and other biologically disfavored persons, who might otherwise be deemed not worthwhile. The concern is that such persons would have their lives, security, and other needs "balanced" against other interests, rather than respected as inalienable and non-negotiable rights.[99] Life as a person with disabilities, an abnormal person, or even an average person who refuses to bring his or her body into conformance with current cultural norms, would become a "choice" that would not be protected.[100]

For instance, some argue that treating embryos as mere property could cause the destruction of lesbian and gay persons or persons with disabilities through genetic selection of "straight embryos" or "able-bodied embryos."[101] The "gay embryos" or "disabled embryos" could simply be destroyed, as one's property can be destroyed. Or perhaps the "straight, able-bodied embryos" would command a higher price in the

95. Stem cell research is widely discussed for its potential to lead to the development of treatments for diseases, and animal–human hybrids might be used to develop organs for transplantation into humans. *See* Bagley, *supra* note 2, at 472, 520.

96. *See* Kass, *supra* note 88.

97. *E.g.*, Volokh, *supra* note 4, at 1835 (arguing that forbidding organ sales violates the right to self defense of those who need organs).

98. *See* Jean Bethke Elshtain, *The Body and the Quest for Control*, *in* IS HUMAN NATURE OBSOLETE?: GENETICS, BIOENGINEERING, AND THE FUTURE OF THE HUMAN CONDITION 155, 161 (Harold W. Baillie & Timothy K. Casey eds., 2005).

99. *See, e.g.*, *id.* at 155–61.

100. *Id.* at 161 (arguing that parents who chose not to abort children with Down's Syndrome and other disabilities would be viewed as having made a choice to do so that society need not support).

101. *See e.g.*, Jennifer S. Geetter, *Coding for Change: The Power of the Human Genome to Transform the American Health Insurance System*, 28 AM. J.L. & MED. 1, 27 (2002) ("The very phrase 'wrongful birth' suggests that the birth of the disabled child was wrong and should have been prevented." (internal quotation marks omitted) (quoting Taylor v. Kurapati, 600 N.W.2d 670, 688–91 (Mich. Ct. App. 1999))).

market than "gay embryos" or "disabled embryos," creating a significant status harm for lesbian and gay persons and persons with disabilities.[102]

In another example, some egalitarians argue that commodification of the body could permit economic disparities to widen between those who can afford financially valuable upgrades to their bodies and those who cannot.[103] Perhaps an upgrade to the body would give a person more abilities—such as strength or intelligence—that would be a great advantage in the employment market. Those who could not afford such upgrades would find it increasingly difficult to earn a good wage, once forced to compete with those whose greater wealth, or whose parents' greater wealth, afforded them the opportunity to buy into privately owned body modifying technology.

Thus, some members of the dignitarian camp would forbid property rights in embryos, in order to prevent perhaps their destruction and likely their sale on the basis of their genetic content. Others might forbid employers from discriminating against those who fail to get certain upgrades to their bodies.

The idea is that the special status of the natural human body must be maintained to prevent some bodies from being worth less than other bodies; otherwise our impulse is to require "natural" bodies to change, rather than changing the environment and society that is inhospitable to them. In other words, our impulse to accommodate different bodies, especially those that are disabled, would be hindered.[104]

However, intuitive ideas about what is natural and what is not do not necessarily lead to protected bodies because aspects of our environment may feel intuitively natural. For instance, the existence of stairs in many buildings and lack of wheelchair ramps may seem a natural, given state of affairs. It took years of activism and scholarly writing by disability advocates to interrogate the medical model of disability and instead promote the "social model" of disability, under which we can distinguish between physical impairment and the social conditions, such as the use of stairs rather than ramps, that turn an impairment into a disability.[105] As

---

102. Aside from ignoring that cultural pressure to perform a heterosexual identity is, and in many cultures has been, enough to crush and subordinate lesbian, gay, and bisexual practice and identity, this argument also fails to recognize that some persons would prefer to have gay children or to be gay, and would choose those options. *See generally* Eve Kosofsky Sedgwick, *How to Bring Your Kids up Gay*, 29 SOC. TEXT 18 (1991).

103. *See* HUGHES, *supra* note 45, at 130–31 (describing the Center for Genetics and Society's argument that cloning and the creation of inheritable genetic modifications will lead to such a situation).

104. *E.g.*, Elshtain, *supra* note 98, at 163 ("[S]urroundings in which bodies are situated fades as the body gets enshrined as a kind of messianic project.").

105. *See* Corker & Shakespeare, *supra* note 42, at 3 (describing the shift from the medical model of disability to the social model, in which activists raised awareness of the fact that impairment alone does not cause disability, but rather social and economic conditions overlaying impairment).

another example, a person with a severe allergy to unprocessed nuts, or to locally native plants, could easily be subject to a belief that his or her body is improperly adapted to the "natural" environment, if we simply rely on intuitive ideas about what is natural. And lesbian, gay, and bisexual people have long been accused of having "unnatural" sexual desires because same sex sexual activity does not "naturally" lead to reproduction. Thus, what sexual minorities do with their bodies is often left unprotected by the appeal to intuitions about nature.

And indeed, it is not clear that in all cases changing the environment is necessarily more compassionate than changing bodies. It may sometimes be better to alleviate suffering by changing what we have previously thought to be "the central core of our humanity,"[106] rather than by changing the environment. A formalist approach that reveres the organic body or listens to the "wisdom of repugnance" does not adequately allow for this, and can therefore seriously harm the rights of those with different bodies, too.

For instance, transgender persons who choose to obtain body modifying surgery or engage in body disguising dress practices are changing their bodies, rather than the environment, both of which are probably contributors to gender identity disorder ("GID").[107] It is unclear why the insistence that we change the environment, rather than permitting transgender persons to change their bodies and clothing, is a superior response to GID. This is especially so when we realize that awareness of and respect for the bodies of transgender persons may contribute to positive changes in the cultural environment. Transgender persons who obtain body-modifying surgery or dress in a manner that obscures parts of the body do not necessarily accommodate or reinstantiate social constructions of gender—these practices can destabilize those constructions.[108] Questioning gender constructions and stereotypes can then lead to changes in the environment and social order, rather than stubborn refusals to change the environment in which our bodies exist.

As an example of where this desire to glorify and dignify the natural body can lead, we can look to Jean Elshtain. Elshtain at times appears to be a defender of those who are different, such as persons with disabilities and sexual minorities, but she also expresses concern with a society that might no longer feel disgust towards a human body "riddled with pieces of metal."[109] Yet, that body riddled with metal may be a person with disabilities who, like all of us, is using the available technology to live the

106. Elshtain, *supra* note 100, at 167.

107. I use the term GID because it is a well-recognized term, not to express any agreement with the medical characterization of many transgender persons as having a "disorder." *See* Dean Spade, *Resisting Medicine, Re/modeling Gender*, 18 BERKELEY WOMEN'S L.J. 15, 19–21 (2003).

108. Judith Butler, *Imitation and Gender Insubordination*, *in* THE LESBIAN AND GAY STUDIES READER 307, 313–14, 318 (Henry Abelove et al. eds., 1993).

109. Elshtain, *supra* note 100, at 167.

best life possible. Alternately, it might be a person who elects to have numerous piercings as an important religious or identity related practice. Revulsion towards such a body doesn't seem compassionate at all.

I believe the worth of different bodies, and the rights of persons with different bodies, is better located in the public's interest in people with different bodies as a means of facilitating cultural change, rather than in a naturalist or intuitionist account of "what is a human." For instance, we can value the bodies of persons with hearing impairments by looking at the richness and beauty of sign language, rather than by valorizing the "natural" state of the impairment and rejecting the wishes of the deaf child who desires a cochlear implant.[110] This way of valuing different bodies—based on their important place in cultural evolution—might work better for persons with disabilities because it recognizes the fact that people are constituted by and within technology. It is persons with disabilities for whom this fact makes a great deal of material difference. In contrast, the naturalist approach might lead to less support for welfare rights to technology that would assist persons with disabilities. Why would a government-funded health care system pay for a cochlear implant, or for a wheelchair, if deafness or inability to walk were revered as "natural"?

The naturalist account is of course similarly dangerous for persons with minority sexual identities. What is "naturally human" or "intuitively" beautiful or repulsive is always contested, and plenty of persons who promote this formalist approach also promote the subordination of sexual minorities and women.[111] If we reflect on how societies have historically treated sexual minorities and persons with disabilities, the "wisdom of repugnance" does not, in my view, seem to be the safest approach for these groups.

Abandoning a naturalist or intuitive source of rights in favor of this approach, geared at social and cultural change, need not mean that people's bodies are only worth something when they have the capacity for rational, political engagement. Instead, we can recognize that bodies are sites of identity and cultural performance, and are therefore worth something to others, even when they are mentally or otherwise disabled in a way that makes traditional political participation impossible. People can be loved, be part of society, and be part of our culture without being

110. In the documentary film *Sound and Fury*, a young child's deaf parents forbade her from receiving a cochlear implant, after much debate with other members of the deaf community and both hearing-impaired and non-hearing-impaired family members. They determined that the implant would be a rejection of the rich and valuable deaf community they were a part of and that was often discriminated against unjustly. SOUND AND FURY (PBS 2000). While views on the parents' decision may vary, there is something tragic in the circumstances that led to the parents being unable to value both the deaf community and their daughter's desire to experience what hearing is like.

111. *E.g.*, Leon R. Kass, *The End of Courtship* (pt. 1), BOUNDLESS, Oct. 13, 2005, http://www.boundless.org/2005/articles/a0001154.cfm (arguing that women should be modest and mothers).

part of our politics.[112] Thus, we don't need an organic account of who is naturally a human in order to value many different kinds of bodies. We can recognize the importance of embodiment without revering the types of embodiment we are used to, or responding with disgust to the types that we are not.

B. What's Normatively Wrong with Autonomy

I have argued that the property-like approach to bodily integrity is descriptively wrong about property—fundamental property rights are not violated merely because they are heavily regulated. But as noted in Part II, some might argue that normatively, we *should* protect either property rights in the body or even all property rights against regulation, as a component of autonomy. In this section, I argue against the use of autonomy as a basis for property-like fundamental rights in the body or any other form of property that would trump typical political concerns such as public health or even parentalism.

One might argue from an autonomy or personhood standpoint that we *should* carve out a property-like right in our bodies, marked off by the physical borders of the body. This right would be different from other property rights in that it would be freer from restraints on alienation and would truly serve as a trump of typical political concerns driving regulation. The idea is that the body, more than property, is particularly fundamental to personhood or autonomy.[113]

But property, too, can be a crucial part of personhood and autonomy, much more so than certain body parts. Margaret Radin, for instance, has persuasively argued that certain forms of property are properly understood as personal in nature, such as wedding rings or houses.[114] The clothes I choose to wear may be a more important component of my personhood than my fingernails, as another example.

Regulation of property can even impact personhood via the body itself. This occurs whenever property is tied to the use and experience of one's body. For instance, if the state forbids the sale of corsets, makeup, and high heels, it infringes on the manner in which I make my body culturally visible and the manner in which I physically experience my body. The same is true when the state regulates sex toys and medical devices. Food and drug regulation limits what we can ingest, and the Controlled

112. *See* MARTHA C. NUSSBAUM, FRONTIERS OF JUSTICE: DISABILITY, NATIONALITY, SPECIES MEMBERSHIP 96–106 (2006).

113. *Cf.* Jed Rubenfeld, *The Right of Privacy*, 102 HARV. L. REV. 737, 800 (1989) (arguing that control of the body is "formative").

114. *See* Radin, *Property and Personhood*, *supra* note 20, at 959–61.

Substances Act can determine whether someone's body exists in pain or not.[115]

By arguing that the body is not a special repository of autonomy or personhood, I do not deny that the body may have a special *relationship* to autonomy, stemming from the cognitive force that being physically controlled can have. It may be impossible in many circumstances to mentally "overcome" experiences like pain or physical restriction, and therefore control of another's body may amount to the destruction of that person's autonomy[116] in a way that control of their property could not compare. Torture has, for this reason, been described as "unmaking the world" of the victim.[117] It is easy to see, with this in mind, why one concerned with autonomy would argue that the government should not be permitted to control the individual's body in a manner that is cognitively overpowering, such as by forcing the individual to endure pain, requiring the individual to experience pregnancy (or prohibiting the individual from having such an experience), or prohibiting the individual's use of psychoactive drugs.

But the claim that regulation of kidney sales, prostitution, manual labor, and other forms of body commodification interferes particularly egregiously with autonomy merely because these activities involve parts or uses of the body would not be supported. An environmental regulation that prevents me from chopping down trees on my land and selling them is no less cognitively overpowering than a public health regulation that prevents me from selling my kidney or engaging in prostitution. Even though having a kidney removed or engaging in sex might have cognitive effects that an individual desires and that cannot be duplicated in any other way, a prohibition on engaging in these activities for money doesn't prevent the individual from having those experiences at all: the individual can still give the kidney away for free, or consent to sex for free.

One might respond to this with the even stronger libertarian view that all property rights should trump typical regulation, whether property rights in the body or not. The problem with this view normatively is that commodification itself, whether of the body or other material, is not en-

---

115. *See* Gonzales v. Raich, 545 U.S. 1, 7 (2005) (noting, in the context of a case upholding the Controlled Substance Act's prohibition on possession of marijuana as a valid exercise of Congress's commerce clause authority, that plaintiff's "physician believes that forgoing cannabis treatments would certainly cause [her] excruciating pain and could very well prove fatal").

116. Rubenfeld, *supra* note 113, at 788–89:

> Yet power need not be directed at the undeveloped mind to have this effect; it may also do so if directed at the fully-developed body. . . . Indeed, bodily control may be the more effective medium to the extent that thought cannot, as it were, meet such control head on, as it might when confronted by an idea that it is told to accept. . . . [I]ts effect can be *formative*, shaping identity at a point where intellectual resistance cannot meet it.

117. *See* ELAINE SCARRY, THE BODY IN PAIN: THE MAKING AND UNMAKING OF THE WORLD 37–38 (1985).

tirely unproblematic for autonomy, as choices can be coerced not only by the state but also by private actors. If the state attempts to promote autonomy by refraining from all regulation of what can and cannot be done with the body altogether, those with private sources of power will be free to interfere with the choices of others. When we live as members of a society, our choices concerning our bodies will never be absolutely "free," in the sense that they are uncoerced, uncontrolled, or undetermined by others.[118]

For instance, when a person has a disability and obtains prosthetics to assist with the disability, this is in part because the person's body has been socially constructed as disabled. If we all used sign language to communicate, fewer deaf persons would desire cochlear implants. In another example, if we failed to classify persons on the basis of the shape of their genitals, fewer transgender persons would desire genital surgery. If actresses didn't receive massive sums of money for looking eternally young, would so many of them "choose" to get injected with Botox?

Someone with a great deal of economic power can offer money to an employee in exchange for the performance of a job, such as manual labor, that eventually damages the employee's body.[119] Or, someone who owns many buildings will, through her choices about whether to install wheelchair ramps, influence the mobility of many persons with disabilities. Markets that reward attractiveness can incentivize cosmetic surgery.[120] Even the most typical forms of employment, such as waitressing or bartending in a "uniform of makeup,"[121] raise concerns about whether selling control over one's body in this way is a choice made "freely."[122] And does someone have sufficient autonomy if he desires a wheelchair for basic mobility, but cannot afford one and is not provided one by his society?[123] In other words, persons are unavoidably embodied in the context of their environment, an environment largely made up of legal relationships of property and contract.

---

118. *See* MICHEL FOUCAULT, DISCIPLINE AND PUNISH: THE BIRTH OF THE PRISON 209–222 (Alan Sheridan trans., Vintage Books 2d ed. 1995) (1977); *see also* Rubenfeld, *supra* note 113, at 770–81 (1976) (making this point with respect to sexuality and Citing Foucault's *The History of Sexuality*).

119. *See supra* text accompanying notes 36–39.

120. *See supra* note 42 and accompanying text.

121. Jespersen v. Harrah's Operating Co., 392 F.3d 1076, 1084 (9th Cir. 2004), *aff'd*, 444 F.3d 1104 (9th Cir. 2006) (en banc).

122. *See supra* note 40 and accompanying text.

123. *See* Convention on the Rights of Persons with Disabilities, G.A. Res. 61/106, arts. IV, XX U.N. Doc. A/RES/61/106 (Jan. 24, 2007) (stating obligations of parties to promote research into and affordable access to aids for persons with disabilities, including mobility aids); *cf.* Dean Spade, *Documenting Gender*, 59 HASTINGS L.J. 731, 782–91 (2008) (documenting the double bind created by, on the one hand, the exclusion of gender confirming medical care such as sex-reassignment surgery or hormone therapy from Medicaid and other state funded health coverage, and, on the other hand, state demands that transgender persons obtain such treatment in order to legitimize their status and be appropriately gender classified).

In critiquing the notion that inviolable private property-like rights best promote autonomy, I am not advocating the eradication of private property. That, too would threaten autonomy, since state coercion is no less a form of coercion than economic coercion. Moreover, even without economic and state coercion, social and cultural coercion remain. For instance, one might obtain cosmetic surgery even in the absence of financial incentives to do so because one desires an intimate relationship, and cultural norms may prefer people with certain kinds of bodies. Or, one may be pressured by family members to donate a kidney to a sibling.

Thus, there is no mechanism by which we can "maximize" autonomy over the body. All we can do is make decisions about where to allocate the coercive power that will inevitably influence individual choices concerning the body.[124] We might choose to allocate all that power in totally unregulated private hands, through a strict libertarian "no state action" doctrine. Or, we might choose to allocate that power in the hands of those with cultural capital, while forbidding or limiting its exercise in economic transactions. Thus, we might limit rights to contract away one's control over one's body, while permitting families and other social groups to coerce individuals into giving up that control. The problem with the autonomy argument is that it doesn't help us choose from these alternatives. We cannot compare all allocations for "levels of autonomy." Some would even say that the notion of autonomy is itself essentially an illusion for anyone who lives in a social context.[125] But even if we do not take this strong a position in critiquing autonomy, it is still a poor basis for a right in one's body that trumps typical regulation.

C. The Capability Approach—Combining Dignitary Approaches with Liberal "Choice"

There are some essentially liberal justifications of human or fundamental rights that acknowledge the problem of private coercion. These accounts therefore reject the libertarian approach in favor of one that might require positive efforts to promote truer autonomy. In this section, I consider and reject a strong example of such an account as an alternate way of justifying human rights in the body.

Martha Nussbaum's capability approach, drawing on Amartya Sen's capability approach to measuring welfare,[126] uses an intuitive approach to reach a list of "capabilities" she argues are essential to human dignity,[127] and she includes bodily integrity in this list.[128] However, she

---

124. *See* FOUCAULT, *supra* note 118, at 194.

125. *Id.*; *see also* MICHEL FOUCAULT, THE HISTORY OF SEXUALITY 159 (Robert Hurley trans., Pantheon Books 1978) (1976) ("The irony of this deployment is in having us believe that our 'liberation' is in the balance.").

126. *See generally* AMARTYA SEN, DEVELOPMENT AS FREEDOM 131–37 (1999).

127. NUSSBAUM, *supra* note 1, at 74–75.

128. *Id.* at 78.

would not require one to maintain one's own bodily integrity, but would only require, in a liberal way, that each person has the *capability* to do so.[129] Moreover, she includes holding property on the list of capabilities.[130] Thus, her approach to rights over one's body combines elements of the traditional liberal autonomy idea with the intuitionist dignity idea.[131]

However, what distinguishes Nussbaum from a libertarian is that for this capability to be provided to all, in the face of the problem of private coercion and control described above, positive government action may be necessary. As a result, the capability for bodily integrity will in many cases need to come at the expense of some portion of the capability to hold property or obtain employment. It wouldn't make sense to "trade off" these two important and essential capabilities against each other, under the approach,[132] but it is also true that not all the capabilities can necessarily be maximized simultaneously. They are instead treated as minimum thresholds—thresholds that incidentally do not appear to be met for many of the world's citizens. Thus, one capability might be reduced in order to promote another, as long as one did not reduce the capability below the minimum threshold. This approach is intended to be flexible,[133] which is part of why it is very coherent, as well as compassionate and realistic in a way that neither the "wisdom of repugnance" approach nor the libertarian approach is. However, the capabilities approach does not seem to help us answer the question of *when* freedom of contract or property must be regulated in order to provide and promote the capability for bodily integrity, and when it should not be. The purpose simply doesn't seem to be to provide a principle for deciding that question.

Moreover, the justification for regulating contract, property, and other forms of individual choice stems, at bottom, from intuitions. Nussbaum's approach is self-consciously "intuitionist."[134] Thus, despite her use of "capabilities," rather than fixed requirements for a life worth living, this soft liberal approach still takes admittedly biased intuitions and uses them universally[135] as a justification for regulating behavior. Nuss-

129. *See id.* at 87.

130. *Id.* at 80.

131. *Id.* at 91.

132. *See* NUSSBAUM, *supra* note 79, at 167.

133. *Id.* at 78–79:

> I consider the list as open-ended and subject to ongoing revision and rethinking, in the way that any society's account of its most fundamental entitlements is always subject to supplementation (or deletion).
>
> I also insist . . . that the items . . . be specified in a somewhat abstract and general way, precisely in order to leave room for the activities of specifying and deliberating by citizens and their legislatures and courts.

134. *Id.* at 174–75.

135. *Id.* at 78 ("The capabilities approach is fully universal . . . . The approach is in this way similar to the international human rights approach . . . ."). It is important to note, though, that Nuss-

baum is laudably cautious about all this, and indicates for this very reason that the list of capabilities must be updated periodically, but the strong risk of biased intuitions being used to justify oppressive regulation still remains.

IV. WHAT THE BODY HAS TO DO WITH FUNDAMENTAL RIGHTS

Does all this critique mean that human rights have nothing to do with bodies? Do rights have nothing to do with controversial questions such as whether to permit prostitution, whether the state may force vaccination, or whether organ selling should be permitted? In other words, should all these questions only be resolved on the basis of typical political and welfare concerns, with no rights "trumping" those concerns? Having critiqued the autonomy, dignity, and capability approaches to resolving these questions, in this Part I propose a different ground for rights in bodies. However, this justification would not support a right to "bodily integrity," alienable or not.

A. Proceduralist Theories of Rights

Autonomy, dignity, and even compelling combinations of such concepts are not the only reasons in which we can ground fundamental rights. Rights can also be vehicles for promoting social and cultural change—ways of helping us update our intuitions. Free speech rights, for instance, have been promoted on this basis.[136] While laws, social norms, and cultural values will always determine what choices are available to us,[137] we can use rights as vehicles for ensuring that those norms can change, and that their evolution is not determined solely by orthodoxy.[138] Such a conception of rights would promote speech rights, for instance, on the proceduralist theory that this contributes to a democratic culture, or even to a search for truth,[139] rather than on an autonomy theory. This might lead to obligations that the government subsidize the speech of those with less capital, or limit freedom of contract so as to protect employee and tenant speech as against the property rights of employers and landlords.

---

baum does not argue that the universal nature of the list "license[s] intervention with the affairs of a state that does not recognize them." *Id.* at 80. She states that "military and economic sanctions are justified only in certain very grave circumstances." *Id.*

136. *See* ALEXANDER MEIKLEJOHN, FREE SPEECH AND ITS RELATION TO SELF-GOVERNMENT 26 (1948) (arguing that access to unconstrained information and ideas is necessary to an informed populace and the functioning of democracy).

137. *See* Kahan, *supra* note 14, at 115 (arguing that even when consciously justified in secular, pluralist terms, laws will often be driven by the expression of cultural values and repression of deviant moral values).

138. Ramachandran, *supra* note 9, at 18.

139. D.A. LLOYD THOMAS, IN DEFENCE OF LIBERALISM 36 (1988) ("The case for a liberal set of individual rights does not rest on the assumption that we already know what is intrinsically valuable. Rather, it rests on a plausible claim about how it is possible to have better formed beliefs about what is valuable.").

Of course, law cannot eradicate social norms, and we probably wouldn't want it to. However, it is important that those norms be capable of changing. In order for culture to be capable of change, individuals and subcultures must be better empowered to challenge and contribute to the construction of culture: some level of cultural velocity must be present.[140]

The goal of maintaining cultural velocity shares similarities with Jack Balkin's thesis that in addition to a democratic political system, we need a democratic culture.[141] Protecting the ability of individuals to form affiliations, beliefs, and values—to engage in making culture—is important, I argue, not because those abilities are crucial to autonomy, but rather because those abilities contribute to cultural exchange and evolution.[142]

B. Cultural Velocity and Identity

What I add to this conception is the view that identity is a precursor to cultural and political meaning making. I define identity as the particular values, beliefs, and aspects of ourselves that we deem so important we consider them self-defining. "Our aversions, desires, beliefs, and choices all make up our identity, but our identity in turn then affects our aversions, desires, beliefs, and choices."[143] "Even when an aspect of identity seems 'unchosen,' such as a biological sex or an ethnicity, we still choose, albeit sometimes within very strong and other times within very weak constraints, whether that 'immutable' trait will be part of our identity."[144] Although those choices will always occur within constraints, law can carve out some space for the exercise of agency in the construction of identity. In this way, identity can be self-defining in the good sense of a set of values and practices that one holds dear (but nevertheless could be different), rather than self-defining in the bad sense of a set of stereotypes about a group that are resistant to change. Why would we

---

140. The idea of cultural velocity as a ground for deriving countermajoritarian rights shares similarities with a consequentialist argument made by David Lloyd Thomas in *In Defence of Liberalism*. *See id.* Thomas defends liberalism on utilitarian grounds, but on a unique form of utilitarianism that he calls "experimental" utilitarianism. He rejects traditional utilitarianism because it would require defining what is of utility—what is good, or of value, and he argues that we do not know the answer to that question. However, he argues that we can defend those liberal rights, which would contribute to discovering what is of value, by permitting experiments, in a sense, in what should be valued. *Id.* at 36–37.

In a future piece, I will more thoroughly explore the concept of cultural velocity as a justification for human rights more generally. For purposes of this article, however, it is sufficient to note that such a concept could explain or justify rights in bodies, but not rights over the "integrity" of bodies.

141. *See* Jack M. Balkin, *Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society*, 79 N.Y.U. L. REV. 1, 3 (2004).

142. *See* Ramachandran, *supra* note 9, at 31–33.

143. *Id.* at 32.

144. *Id.* For a similar conception of identity as "choice" rather than as "discovered" fact, see AMARTYA SEN, THE ARGUMENTATIVE INDIAN: WRITINGS ON INDIAN HISTORY, CULTURE AND IDENTITY 350–52 (2005).

want to promote diversity and reformability of identity? Because identity is determined in part by cultural norms, but it is also a ground from which to rally for changing those norms.[145]

We cannot ensure total "freedom of identity" or "freedom of personhood" for the same reasons we cannot maximize "autonomy." Such freedom is to some extent an illusion, given that even our physical bodies are constructed in part by their social context. Our choices are heavily constrained. But in the face of those constraints, we often see resistance. Sexual minorities, racial minorities, disability activists, and those excluded from aristocracy and nobility are just a few examples of persons who have engaged in culture wars, perhaps not securing themselves "liberation,"[146] but certainly moving culture in radical ways. This is one of the reasons often given for protecting speech rights—such rights facilitate that cultural and political change.

What is often neglected, though, is that resistance to current cultural norms need not take the form of traditional political activism, speech, and interest group formation. Certain identity practices, such as dress, sexual practice, speech, mannerism, body modification, and the like, can themselves be subversive, destabilizing acts that have cultural effects.[147] Judith Butler has described drag as such a practice, calling it "gender insubordination."[148] These culturally disrupting, identity disrupting practices have been valorized by postmodernists and queer theorists, though their work to do so has often seemed too "academic," not having any practical import.[149] Part of what makes this celebration of subversion seem pointless is that the prescription to disrupt identity or subvert cultural norms has limited use for someone who will lose her job, be kicked out of school, or sent to jail for doing so.[150] That is where legal rights come into play. Law can carve out some space for individuals, subcultures, families, and other groups to form different, challenging identities, and even reform them, yet still have a job, shelter, and other needs met that would permit participation with the broader culture. In this way, rights can be used to ensure that some level of cultural velocity is maintained.

---

145. Ramachandran, *supra* note 9, at 93.

146. *See* FOUCAULT, *supra* note 118, at 289.

147. *Cf.* Judith Butler, *"Appearances Aside,"* 88 CAL. L. REV. 55, 63 (2000).

148. *See* Ramachandran, *supra* note 9, at 22 (citing Butler, *supra* note 108, at 307).

149. As Seidman has explained in an illuminating history of queer theory's development, "[t]o the extent . . . that poststructuralist perspectives . . . [has become] a politics of the disruptive gesture, they lack coherence." STEVEN SEIDMAN, DIFFERENCE TROUBLES: QUEERING SOCIAL THEORY AND SEXUAL POLITICS 136 (1997). As he has also noted, "Underlying this politics of subversion is a vague notion that this will encourage new, affirmative forms of personal and social life, although poststructuralists are reluctant to name their social vision." *Id.* at 134.

150. *See* Butler, *supra* note 147, at 63:

> For this challenge to take place, it must be possible for a person whose appearance calls the category of the person into question to enter into the field of appearance precisely as a person. . . . [A] power that is 'had' to the extent that such a person is not first defeated by the powers of discrimination.

C. Applying Proceduralist Rights in Practice

Of course, many behaviors are a part of identity formation and reformation. Dressing a particular way is part of identity formation and reformation, but so is smoking, reading certain magazines, speaking in particular dialects, and so on.[151] Law cannot protect the freedom to engage in all such behaviors, especially if we want law to protect these behaviors not just against the state, but to also positively regulate employers, landlords, and other private actors who would use their economic power to direct these behaviors. Thus, although the proceduralist principle of promoting cultural velocity represents a non-intuitionist, consequentialist ground for recognizing certain rights, a practical application of this principle to create actual legal rights must depend on cultural context.[152] We will have to choose some identity formative behaviors to protect legally.

In many societies, freedom of speech will be a sensible right to carve out as a means of promoting cultural velocity. I have argued that freedom of dress in the United States is also a sensible right to carve out, in part because of the unique way in which dress is experienced as both social and deeply personal, experiences that are historically and culturally contingent.[153] I argued that it even ought to be protected to some degree against private regulation, such as in the workplace. How that protection could be functional in the context of American capitalism of course also required attention to the cultural and political context. Thus, a relatively weak "reasonable accommodation" framework was what I proposed.[154] But in a society in which physical appearance is not a particularly common site of identity exploration and formation, a site that is frequently experienced as uniquely personal or physical in nature, for example, it might not make sense to recognize a freedom of dress. Or, in particular professions, such as modeling and acting, it might also not make sense to recognize a freedom of dress.[155]

Once rights that promote cultural velocity are fleshed out in this way, with an eye to the particular context, one might question how this differs from many modern forms of liberalism, such as Martha Nuss-

151. *See* Ramachandran, *supra* note 9, at 25, 33.

152. *Cf.* Angela R. Riley, *Good (Native) Governance*, 107 COLUM. L. REV. 1049, 1078 (2007). Riley has argued that good governance need not look like liberal democracy in every cultural context, but has still articulated some core components of good governance that are broadly applicable to many sovereigns, such as the ability to dissent and the ability to exit. *Id.* at 1061–80. The cultural velocity principle would likely result, in virtually every society it were applied to, in protection of the formal ability to dissent and to exit. I believe, however, that these formal abilities would rarely be quite enough to ensure a significant amount of cultural velocity—the amount required to ensure that a culture can evolve and meet new challenges. Room to develop the kinds of different identities and affiliations that lead to the desire to dissent and exit is required, in my view.

153. *See* Ramachandran, *supra* note 9, at 25.

154. *Id.* at 61.

155. In fact, I proposed a categorical statutory exception for these professions. *Id.* at 63.

baum's intuitionist concept of capabilities. Her approach also varies based on the particular context, and it includes capacities for mental and social development that might produce very similar results to the cultural velocity approach's concern with identity formation and reformation.[156] One might argue that by using culturally contingent facts, such as the importance of appearance manipulation to identity formation in the United States, I have imported what amounts to an intuitionist approach through the back door.

The important difference is the grounds from which I draw the need to protect identity formation. I draw not from intuitionist or dignitarian grounds, nor from an autonomy ground, but from the consequentialist ground that cultural velocity is of procedural value in all cultures because it helps secure the ability to adapt, whatever one's view of a good life is. This means that claims about what will best promote cultural velocity are *empirically falsifiable*. If I argue for a legal protection of dress in the form of a right based on the culturally contingent fact that dress is an important site of identity formation, and dress loses this cultural salience, I would have to change my conclusion.[157] If it turns out that protecting dress in the form of a right hinders, rather than promotes, cultural velocity, I would also have to change my conclusion. What rights we protect through an application of the cultural velocity approach will therefore change when cultural facts change, when scientific understandings change, and when technological advancements change.

Intuition, on the other hand, is by definition slow to change. Intuitions might be formed through repeated observations of the world,[158] or through a complex set of cultural meanings,[159] or both. But however they are formed, intuitions are by definition those beliefs that have become "immediate," and involve "knowing or sensing without the use of rational processes."[160] They are therefore resistant to challenges from empirical evidence.[161]

156. NUSSBAUM, *supra* note 1, at 78–79 (including "senses, imagination, and thought" on the list of capabilities, as well as "practical reason").

157. In fact, it is entirely within the realm of possibility that I will be forced to change my conclusion about freedom of dress within my lifetime. Because working from home is becoming more and more common, people are spending less time interacting with each other visually, and more time online. If this trend develops dramatically, the importance of dress both as a means of forming an identity, and as an impetus of cultural velocity, might diminish in the future.

158. *See* Christine Jolls, Cass R. Sunstein & Richard Thaler, *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471, 1477–78 (1998) (describing the use of mental heuristics as an example of bounded rationality).

159. *See* Jerry Kang, *Trojan Horses of Race*, 118 HARV. L. REV. 1489, 1508 (2005) (describing racial schemas, or meanings, as developing not merely in response to data perceived by one who adopts the schema, but also growing out of complicated cultural meanings).

160. AMERICAN HERITAGE DICTIONARY 919 (4th ed. 2000).

161. *See* Jolls et al., *supra* note 158, at 1477–78; Kang, *supra* note 159, at 1508.

D. Cultural Velocity and the Body

A cultural velocity approach seeks to protect diverse formations of identity, but of course many behaviors are related to identity formation. No society can protect all in the form of a legal right. We will have to protect those areas of behavior which are most commonly culturally meaningful, or which are unique in other ways that make them convenient to carve off from other behaviors in the form of a right. Do reasons such as this exist to carve off legal rights protecting control over one's body?

1. The Body Is Political

The body is the site of our coming into being socially, and in the United States, as well as in many societies, appears to be a particularly common site of identity exploration, performance, and formation. I am not arguing that the body is socially or culturally important in any "natural" or "essential" way. The argument here is decidedly not an argument that treats the body as sacred, natural, and immutable.[162] Nor is it an argument that treats the body as distinct from and unaffected by the external world and culture.

Indeed, I am positing that the body is not sacred, is deeply alterable, and that "natural" and "unalterable" attributes of our bodies—such as the color of our skin, the shape of our genitals, the color of our eyes, and the texture of our hairs—may be no more important to many of us than "artificial" and "alterable" aspects of our bodies, such as our tattoos, piercings, jewelry, clothing, dyed hair, braided hair, hip implants, or canes.[163]

Rather than treating the body as a temple, I am recognizing its inescapable role—however historically, culturally, or politically contingent—in our experience of self. I recognize that the body may hold no more inherent, acontextual value than chattel. This is precisely why I argue that we should not create a "right in the body" that rests on a formal separation between persons and property, privileging outdated constructions of a biological or "natural" body.

On the other hand, the body's manipulation and alteration as an identity formative practice has become widespread. The body and the world around it have begun to bleed into each other: extreme body mod-

162. *See* Shell, *supra* note 3, at 333–34, 345–47 (discussing Kant's "concept of human dignity" and bioethics in stem cell research).

163. In fact, the distinction between the natural or immutable and the artificial or choice-based aspects of our bodies is a slippery one. Skin color, genital shape, eye color, and hair texture are all alterable, with varying levels of effort. Body weight is alterable, but sometimes only with great effort and social support. Tattoos are close to unremovable. Hip implants are removable, but only at great medical cost to those who have them, and once implanted, they are internal to the owner. Are scars from accidents "natural," or "artificial"? What about scars that are deliberately obtained as a form of cosmetic body modification? Scars that result from surgery? Does it matter if the surgery was defined as "elective" or "medically necessary"? Whose definition for those categories should we use?

ification is on a more visible rise, clothes and other objects seem more and more like parts of our bodies and extensions of ourselves, transsexual persons' alterations of their gender are more and more visible, some Asian Americans have opted for eyelid surgery, and Michael Jackson is widely thought to have changed his skin color. And yet, all these trends don't seem to represent a rejection of the body as having any importance at all. These are not even trends toward the soul and mind taking precedence over the body. If transsexual identity were about transcending the importance of genital shapes, then why would so many transsexuals endure the oppression they endure, spend the money they spend, and enter into complicated relationships with the medical establishment to change that which doesn't matter—the shape of one's genitals? Rather, these manipulations, alterations, adornments, and extensions of our bodies themselves carry a great deal of meaning to most of the people engaged in them.[164]

This need not be the case in some science fiction world of the future, and it need not be universal. However, the fact that the meaning of these body manipulations is contingent on a social and cultural context makes them no less foundational a component of identity. It does, however, raise the question of whether the natural body actually makes sense as a repository of special legal status. If we are accounting for the frequency with which the body appears to play some especially important role in identity formative practices, mapping the borders of a legal right precisely onto the physical borders of the "natural," human body does not make sense; for it is often "artificial" and "elective" manipulations and adornments to the body that are identity performative.

There are many activities that are commonly experienced in terms of their physicality, activities which are experienced as manipulations of the body, but which do not involve alteration of the natural, biological, integrated body. For instance, American feminists who argued for the right of women to wear "sensible clothes," at the same time they argued for the rights of women to vote and have jobs, did so not just because the clothes that women were coerced to wear held a political meaning that subordinated women, but also because the clothes were physically restrictive. Skirts and petticoats were experienced as a *physical* restraint on women.[165] On the other hand, those who militate for freedom of speech have rarely done so in terms of the right to control the air coming in and out of their lungs and the movements they make with their mouths, de-

164. Ramachandran, *supra* note 9, at 39.

165. *See* KARLYNE ANSPACH, THE WHY OF FASHION 329–30 (1967) (describing the Bloomerism movement to replace skirts with pants); CHARLOTTE PERKINS GILMAN, *The Force Called Fashion*, *in* THE DRESS OF WOMEN: A CRITICAL INTRODUCTION TO THE SYMBOLISM AND SOCIOLOGY OF CLOTHING 107, 116–17, (Michael R. Hill & Mary Jo Deegan eds., 2002); CHARLOTTE PERKINS GILMAN, *Hope and Comfort*, *in* THE DRESS OF WOMEN, *supra*, at 131, 133–41 (discussing the oppressive nature of fashion).

spite the fact that these behaviors arguably involve the body more intimately in a physical sense, by making use of its internal organs.

Similarly, contracting oneself out for arduous physical labor is often experienced as contracting out one's physical self, while contracting oneself out as a professor—while involving physical activity such as standing and talking—is rarely experienced in this manner.[166] If the reason we should protect rights in our bodies stems from their especially important role in identity formation, in contributing to cultural velocity, then we should focus in particular on those uses of the body which, as a cultural and social matter, are actually experienced as bodily practices of identity formation or reformation.

2. The Body is Personal

As the previous section reminded us, the importance of the body as a site of identity experimentation in many societies cannot be overstated. Many persons consider the ways in which they fashion, adorn, and manipulate or modify their bodies (or refuse to do all these things) to be extremely important to their identity and sense of self. The relationship of these behaviors to physical experience is part of what makes them unique among all sorts of identity performative behaviors.

This uniqueness may make many of these behaviors good candidates for special legal protection. Because not all identity performative behaviors can be singled out, it may make sense for us to at least include those which have the unique aspect of being experienced by many as both personal and political, straddling the line between our separation from and connection to others. But is there any reason to single out the human body *per se*, as opposed to some of its extensions, such as clothing and perhaps other chattel, from other arenas of identity exploration? In this section, I consider whether the fact that we have embodied subjectivities[167] provides such a reason.

Embodied subjectivity describes the notion that we are not cognitively separate from our bodies, that we are not minds or souls simply "inhabiting" or "inside" a physical body. We experience the world not as consciousnesses separate from and encapsulated within a body, but rather *through* our bodies.[168]

For instance, what if I purchased artificial eyes to improve my vision, but those artificial eyes contained a filter, blocking me from seeing images that were deemed to undermine the interests of the company that sold me the eyes? This physical change would alter my subjectivity it-

166. *See* Nussbaum, *supra* note 13, at 693–94 (pointing out that professorial labor, like prostitution, involves contracting out the body).

167. *See, e.g.*, MARGARET A. MCLAREN, FEMINISM, FOUCAULT, AND EMBODIED SUBJECTIVITY 2–3, 14, 83–91 (2002).

168. *See, e.g., id.* at 83–84.

self—my perception and experience of the world. The power to force or even encourage most persons to implant such a filter would be a great power indeed.

The concept of embodied subjectivity shares something with the view that, "rather than speaking of rights in our bodies[,] it would be more appropriate to say that for Hegel we have rights through our bodies . . . an assault on one's body is not experienced as damage to one's property but rather as direct injury to one's self as a person."[169]

Indeed, our brains and nervous systems, which in modern times we often imagine to be our "consciousnesses," are part of our bodies, and thoroughly embedded in our bodies. Nobody has yet succeeded in transferring the contents of a person's consciousness to a location outside the person's body. There is currently no way to "upload" your brain to a computer. Even if there were, it is unclear if the computer's inability to taste, touch, smell, and see would proceed to alter the consciousness in some significant way.[170] Another way to put this "is that all cognitive experience involves the knower in a personal way, rooted in his biological structure."[171] "We have a subjective experience of our own thought processes, but at best only an imagined representation of what goes on in others' subjective experience."[172]

The concept of embodied subjectivity implies that if I change my body, I change the structure through which I perceive and experience—I change my subjectivity. Were we to consider only this, we might promote a human right in the body that maps directly onto the physical, human body. Jed Rubenfeld, for instance, has argued that bodily control should be part of the fundamental right to privacy because it may be a particularly effective way for the state to influence subjects at the "formative" level.[173]

---

169. Uffe J. Jensen, *Property, Rights, and the Body: The Danish Context—A Democratic Ethics or Recourse to Abstract Right?*, *in* OWNERSHIP OF THE HUMAN BODY, *supra* note 37, at 173, 180–81 (alteration in original) (quoting HARRY BROD, HEGEL'S PHILOSOPHY OF POLITICS 69 (1992)).

170. This is not to say that the uploaded consciousness's difference from us would mean it should be treated differently. That question goes to whether legal personhood should be granted to so-called artificial intelligences. *See* Lawrence B. Solum, *Legal Personhood for Artificial Intelligences*, 70 N.C. L. REV. 1231, 1231–32 (1992) (exploring this question). The point is simply that current persons have embodied subjectivities, and we should therefore consider those embodied subjectivities when we think about what rights to grant persons, and how best to fulfill them. It may be that disembodied subjectivities wouldn't need the same rights that embodied subjectivities have. This potentially unequal treatment should not trouble us, as this would simply be a case of treating unlike persons differently, rather than treating like persons differently.

171. HUMBERTO R. MATURANA & FRANCISCO J. VARELA, THE TREE OF KNOWLEDGE: THE BIOLOGICAL ROOTS OF HUMAN UNDERSTANDING 18 (rev. ed. 1998).

172. TERRENCE W. DEACON, THE SYMBOLIC SPECIES: THE CO-EVOLUTION OF LANGUAGE AND THE BRAIN 424 (1997).

173. Rubenfeld, *supra* note 113, at 800–01.

Imagine that we could have filters implanted into our eyes. What if a filter were created that made it impossible for the recipient to see images deemed dangerous to the government, and the government forced us all to have the filters implanted? In contrast, imagine that the state forces us all to paint our houses in the colors of the incumbent political party. We might propose that because we have embodied subjectivities, there is a good reason to be even more frightened of the former state than the latter. If the state can control how its citizens perceive the world, this creates, we might say, a much greater danger of cultural stagnation than if the state can control whether citizens demonstrate support for the government. At least in the latter case, the citizens could reappropriate the incumbent party's colors for some subversive purpose, or could perhaps shift their dissent to other arenas, like dress or speech.

In other words, the idea of embodied subjectivity might provide a factual, scientifically verifiable reason to distinguish between the body and "personhood property" like wedding rings or houses—a reason to create "privacy" rights in the body. However, we should not forget that the body is not *only* personal, due to its important role in our subjective experience; it is also generally political, because it is often the site at which we take on a social form. Thus, some uses of and changes to the body will be more important to the formation and exploration of identity than others, despite the fact that subjectivity is embodied. Moreover, some changes to and uses of the body are commonly experienced in terms of their physicality, such as appearance manipulation, and others are not, such as speech. This is largely a product of social circumstance. Finally, and most importantly, even the subjective, lived experience of the body is inevitably determined in part by social circumstances and coercion. The law *cannot* protect embodied subjectivity by carving off the human body because even this does not make it fully secure from state and private coercion.

Suppose, for instance, that a technological improvement were invented that provided purchasers with additional sixth—or even tenth—"senses." In fact, implants have already been created that would permit the recipient to sense electromagnetic fields.[174] Implants of this sort are property. They are clearly not part of the organic, human body. But the power of the implant seller to retain certain property rights in the implant may become a power to influence the subjectivity of the implant recipient. Such a power is inevitable in a world in which the body and property exist in relation to each other, in which they bleed into each other.

In other words, if there is a "body" that human rights or constitutional law ought to be protecting, it is not the "human body," as defined

174. Quinn Norton, *A Sixth Sense for a Wired World*, WIRED.COM, June 7, 2006, http://www.wired.com/gadgets/mods/news/2006/06/71087.

to mean an organic, physically continuous being distinct and isolated from the surrounding world, but rather the "posthuman body," defined as constructed by and situated within a social and technological context.[175] Protecting this "posthuman body" can't be done by carving it off for special legal status because it can't be carved off at all. Thus, it must be protected through a combination of negative rights against the state and positive regulation of property and contract. Under this view, we should no longer recognize a fundamental right to "bodily integrity," and fundamental rights should abandon the concern with commodification of human bodies *per se*. However, we should choose rights to protect—both as negative rights and through positive regulation—with a recognition that various uses of and changes to the body are especially important to the formation of diverse identities, and therefore, to the goal of cultural evolution.

In sum, the body is indeed a unique arena of identity reformation because of facts about embodied subjectivity—a subjectivity that is always influenced by the world around the body—and because of facts about the body as an extremely common site of cultural, political, and social identity performance.[176] This provides a fair reason in many societies to give special legal attention to identity development and formation that is related to the body. But it does not provide a sensible reason to create a legal right that maps neatly onto the natural, human body. In the next section, I outline what factors we would appropriately consider, from the cultural velocity perspective, when deciding what kinds of coercion over the body should be carved out for special legal attention.

3. Protecting Bodies from Control

Due to the embodied nature of subjectivity, control of a person's body may in fact become control of that person's very subjectivity, directing the identity, thoughts, and beliefs of the person being controlled. But our fear of this result may not be substantiated in every instance in which the body is commodified, even when this occurs in a context of coercion.[177] For instance, when blood is removed from the body, sensory experience and inputs do not appear to change dramatically in the long term. When someone engages in sexual activity, future experiences may be altered significantly, but the change might be a positive one, rather than a negative one.[178]

---

175. *See generally* PEPPERELL, *The Posthuman Manifesto*, *supra* note 51.

176. It may even be the fact that subjectivities are embodied that has partially led to the frequency with which cultures treat the body as a primary site of identity expression, formation, and reformation.

177. *See* Stephen R. Munzer, *Property as Social Relations*, *in* NEW ESSAYS IN THE LEGAL AND POLITICAL THEORY OF PROPERTY 36, 46–52 (2001) (distinguishing between coercion, exploitation, the use of power, and coercion as a neutral term and a negative term).

178. Of course, if the sexual activity results in pregnancy, these changes to the biological structure are probably severe enough to dramatically change the subjective experience of the person,

Sometimes a bodily intrusion or invasion carries with it a destructive social or cultural meaning that one is subordinate or one's identity worthless. Rape has been described in this manner, for instance. Socially and culturally, bodily invasion and control by another is often experienced as a loss of control over one's identity. On the other hand, those who *consent* to sexual intercourse are not thought to generally experience a loss of self-worth or identity in the way that some rape victims may. We can ask ourselves in these varied situations, in what way does the right to refuse or consent to bodily intrusion and invasion promote the possibility of new and creative exercises of agency in the formation of identity?

An interest of countermajoritarian constitutional or human rights law in individuals' control over their own bodies is the interest in what that control promotes—the capacity for social change as security against stagnation. This explains why we might intrude on the libertarian autonomy to contract away control of the body where it better promotes that public interest. Just as state control of bodies might get in the way of bodies being used to explore and reform cultural identity, too much private control of bodies might, too.

We may legitimately fear that an individual will become unable to form their own opinions, thoughts, make their own decisions, and consider a wide variety of cultural affiliations if another controls their body, even if they consented to that control. Thus, there is something sensible underlying the concern many have with transfers of rights in the body, once we conceive of those rights as connected to cultural and political expression and identity development, rather than merely zones of autonomy. Since subjectivity is embodied and bodies are so important culturally, those transfers may threaten our interest in dispersed, rather than consolidated, control of the formation of subjects and identities. If the formation of subjects is left to the control of the market, we may see societies as static as those in which the formation of subjects is rigidly controlled by the state.

If we cared only about autonomy, we would leave persons to suffer the consequences of their choices. That is what it means to have a choice. And if we cared only about a static notion of respect for the natural body, we would often refuse to permit persons to modify their own bodies. But if we also care about freedom of thought, affiliation, and cultural production because it is good for society, then we might want to prohibit someone from entering into contracts that will inhibit those freedoms, in order to prevent those with greater capital from obtaining a monopoly on these crucial contributors to social, political, and cultural change. The goal of a

---

even aside from changes in the way the person is treated socially. Thus, we might want to do our best to ensure pregnancy is a state entered into with some forethought.

society not overly constrained by orthodoxy may require that we forbid persons from entering into too many self-constraints, but would still require respect for most of the choices persons make.

However, getting rid of all social coercion over our bodies is impossible. The relevance of our bodies is largely within a social context. When a property owner places stairs at the front of a building, she makes the bodies of those persons who cannot climb stairs relevant to their experience of the world in a way it would not be had she made the entrance level with the sidewalk. When an employer requires his cocktail waitresses to wear three-inch heels at work, he has not required them, at least in the short term, to modify their natural bodies,[179] but he has changed the way it feels when his employees walk, stand, or even sit. Even "the use of tools and artifacts," such as at work, "requires a degree of incorporation into the body; . . . the 'labor of animation.'"[180] And these uses of the body have lasting effects. For instance, despite the fact that Henry Ford believed that repetitive labor would not "injure[] a man in any way," such injuries do result. (Ford admitted having been told "by parlour experts that repetitive labour is soul—as well as body—destroying, but that has not been the result of our investigations," he argued.)[181]

Not only is all this private coercion inevitable, human flourishing is possible within the context of all this coercion, and is even facilitated by the existence of private repositories of power. For instance, although capitalism is a system that places a great deal of coercive power in private hands, capitalism can provide the freedom to deviate from social and cultural norms that contributes to the creation of new subcultures. These subcultures can then challenge the very norms from which they deviate. For instance, capitalism opened up a venue for the flourishing of queer culture,[182] and gay men and lesbians have found a refuge in commercial venues such as bars and certain sectors of the entertainment industry.

Similarly, although it would seem that the choice to have sex in exchange for money is always, in some sense, coerced by the need or desire for that money, this commodified sexual activity is not necessarily less fulfilling or more corrupted than sexual activity arising out of "love" or other impulses. For instance, in the work of some pro-sex feminists, prostitutes have reported emotional connections and even having "favo-

179. Although in the long term he may have caused them lasting injury.

180. Jain, *supra* note 38, at 32 (quoting Elaine Scarry, *The Merging of Bodies and Artifacts in the Social Contract*, *in* CULTURE ON THE BRINK 85, 97 (Gretchen Bender & Timothy Druckery eds., 1994)).

181. *Id.* at 34 (quoting HENRY FORD WITH SAMUEL CROWTHER, MY LIFE AND WORK 105 (Doubleday, Page & Co. 1923) (1922)).

182. *See* John D'Emilio, *Capitalism and Gay Identity*, *in* THE LESBIAN AND GAY STUDIES READER, *supra* note 108, at 467, 473–75.

rite" clients.[183] These positive stories of prostitution can in turn challenge our conceptions of what sexual activity should be and what role it should play in our lives.

Finally, proponents of cultural property rights for indigenous and other groups, in response to arguments that these rights would "commodify" important elements of a culture, have demonstrated how this commodification need not be destructive to a subordinated subculture, and can be used for progressive ends.[184]

In other words, we should acknowledge that human flourishing can occur in the context of economic transactions, and can even be facilitated by it.[185] In the context of the First Amendment, the United States Supreme Court has already recognized this, holding that forbidding profit from speech is an unconstitutional restriction of speech.[186] Thus, a federal law that prevented civil servants from being paid for speeches and writings has been struck down as overly restrictive of speech.[187] Similarly, one of New York's "Son of Sam" laws was struck down as overly restrictive of speech. The law required convicted criminals profiting from their crimes through books and movie deals to donate the profits to a victim compensation fund.[188] These holdings acknowledge that banning the commodification of a practice, or banning profits from the practice, does not in the context of a generally capitalist economic system "protect" the practice in any sense. Rather, it discourages the practice.

On the other hand, the sometimes deleterious effects of capital on these explorations of and articulations of identity cannot be ignored either. The most visible forms of gay culture have displayed a tendency to idolize the white and the wealthy.[189] And positive stories of prostitutes documented by pro-sex feminists aside, there are plenty of stories of prostitution experienced as exploitative, denigrating, misogynist, and

---

183. Lucas, *supra* note 22, at 252.

184. *See* Carpenter et al., *supra* note 72, at 1026–30; Madhavi Sunder, *Property in Personhood*, *in* RETHINKING COMMODIFICATION, *supra* note 22, at 164, 171 ("[T]here may, in fact, be some room for property in personhood claims if they are grounded on more modern understandings of both culture and property.").

185. *See* Carol M. Rose, *Afterword: Whither Commodification?*, *in* RETHINKING COMMODIFICATION, *supra* note 22, at 402, 404 ("Markets seem inappropriate for some things, but then again, maybe markets are pretty useful for exactly the same things."). Rose also discusses the "market's possibilities for novelty, liberty, and self-fashioning—not to speak of money." *Id.* at 421.

186. *See* Va. State Bd. of Pharmacy v. Va. Citizens Consumer Council, Inc., 425 U.S. 748, 761–62 (1976) (concluding that speech which proposes a commercial transaction is protected by the First Amendment).

187. United States v. Nat'l Treasury Employees Union, 513 U.S. 454, 467 (1995) (holding the government to a high burden to justify its "wholesale deterrent to a broad category of expression by a massive number of potential speakers").

188. Simon & Schuster, Inc. v. Members of the N.Y. State Crime Victims Bd., 502 U.S. 105, 117 (1991) (subjecting Son of Sam law to First Amendment scrutiny because it operated as a "disincentive[] to speak").

189. Ramachandran, *supra* note 9, at 55–56 (using *Will and Grace* and *Queer Eye for the Straight Guy* as examples).

racist. We should acknowledge that systems of capital can facilitate human flourishing but also have an enormous effect on the end results of that flourishing. Leaving that influence unchecked risks the consolidation of power over cultural and social production in the hands of a few. Thus, we may want to soften the effects of capital on practices that promote cultural velocity through regulation of property and contract, rather than through total bans on practices such as commodifying the body.

So what question should we answer when we embark on that project? When should we regulate property and contract in the interest of the special role the body may play in cultural velocity? I propose the following standard: we should ask ourselves, to what degree does the property transfer being contemplated, or the assertion of rights by the property owner, raise the concern that an individual will lose the ability to explore, form, and reform his identity as a social being?

a. Pain

If the control an individual seeks to submit himself to involves pain or extended physical discomfort, the fear may be substantiated. For most persons, subjection to pain and discomfort is cognitively difficult to overcome. Of course, some persons are particularly lucky or resistant; they do not mentally collapse even after years of torture or slavery. But others lose all capacity to interact as social beings. Torture has been described as "unmaking the world" of the victim for this reason.[190] Even those in slavery who are not physically abused can often lose their sense of agency in manufacturing a social identity. The embodied nature of our subjectivities is why the infliction of physical pain through torture can have this effect. In other words, even in the absence of any cultural meaning that torture is disrespectful to the victim, acts like torture may, via control of the body and infliction of pain, too greatly inhibit the individual's development of an identity.[191]

b. Duration

In addition to the degree of pain a particular intrusion on the body entails, the duration of the consequences of subjection to the physical control of another may be relevant to the capacity to form and reform one's own identity. For instance, there is a substantiated risk that someone who sells an entire limb will undergo a change in his or her experience of the world so significant (and permanent) that it may be difficult to cognitively overcome the feeling that one's identity has been altered.[192] Selling oneself into indentured servitude of permanent duration

190. SCARRY, *supra* note 117, at 37–38.

191. *See* Rubenfeld, *supra* note 113, at 776–81.

192. *See* Miho Iwakuma, *The Body as Embodiment: An Investigation of the Body by Merleau-Ponty*, *in* DISABILITY/POSTMODERNITY, *supra* note 42, at 76, 81 (describing the phenomenon of phantom limbs in amputees).

would likely do something similar. But the fear that one will lose the capacity to develop one's identity by selling blood or hair is far less substantiated because neither of these losses is very permanent.

c. Cultural Meaning

Finally, if the control an individual seeks to give up is over the way the body is culturally read and presented, this can implicate law's interest in dispersing control over formation of identities. Thus, when employees contract away to employers their rights to dress as they please, the interest in dispersed control over bodies is implicated. This need not mean that as the right is balanced against other interests, the right always prevails. However, it does mean that positive regulation of such employment contracts might be one way of promoting cultural velocity through individual rights to "bodily" exploration of identity.

In the next Part, I explore a few tentative examples of how we might apply this standard in the United States. These examples will serve to flesh out how the approach might be applied in a particular cultural context. Even when we take into account current social realities about the importance of the body, the legal rights we would then recognize will look different than traditional rights to bodily integrity we have seen before. The examples will necessarily not be comprehensive, and of course, application of the principle will sometimes involve a host of other considerations, especially when positive property and contract regulation is considered. The purpose of the examples, however, is to help elucidate the difference between rights in the human body and rights that acknowledge the body's importance, but stop treating it as an end in itself.

V. EXAMPLES OF THE BODY'S RELATIONSHIP TO FUNDAMENTAL RIGHTS

Applying a human or constitutional right that appropriately protects the public interest in cultural velocity, as it relates to the practical importance of the body in cultural velocity, will necessarily require both negative rights against the state and positive property and contract regulation. Only a combination of both negative rights and positive regulation can help to ensure that the legal and social coercion that bodies are subject to does not lead to consolidated control over culture in the hands of a powerful few.

Some of these example applications will result in very familiar negative rights against the state, ones that both dignitarians and libertarians promote, such as the right to refuse medical treatment or other state attempts to mandate changes to the body. Other familiar rights against the state would include the right to change one's own body, such as the right to obtain tattoos and piercings, to cross-dress, or to purchase medical treatment.

Other applications will consist of protecting rights that increase cultural velocity against private regulation, resulting in property regulation that many dignitarians would agree with, but that strong libertarians would reject. For instance, we might limit the ability to sell one's own body part when that sale would result in long term physical pain and thus cognitively interfere with the ability to develop and reform one's identity.

There will also be applications of cultural velocity to protect rights against private coercion that would be totally unfamiliar to either the dignitarian or libertarian approaches. For instance, my approach might lead us to regulate exchanges in ordinary chattel, such as certain implants and prosthetics. Still other forms of regulation might consist in extensions of employment and housing law, protecting people with different bodies from discrimination in these realms.

I also describe in this Part some examples of property and contract regulation, such as total bans on prostitution and organ selling, that do *not* in fact further properly conceived "rights in the body." This is despite the fact that such regulation is frequently justified on the grounds that it protects against commodification of the body.

A. Rights Against State Intrusion—from Torture to Forced Vaccination

Applying the cultural velocity principle to questions of state intrusion into the body would likely result in the recognition of a relatively familiar negative right against state intrusion into the body. This is because embodied subjectivity creates the danger that many forms of state intrusion on bodies are too cognitively difficult to overcome. If the state can intrude on one's body, and can thereby direct our subjective experiences, this may lead to state suppression of different ways of living, different experiences, and different identities, all of which generally contribute to cultural evolution when left unpunished by criminal law.

Thus, rights to resist severe state imposed pain[193] or very invasive or permanent medical treatment,[194] such as forced sterilization,[195] would likely be included within an application of the cultural velocity approach, just as they would likely be included in the "bodily integrity" rights protected under many dignitarian and autonomy approaches.

193. *Cf.* Baze v. Rees, 128 S. Ct. 1520, 1530 (2008) (noting that what many forbidden punishments under the Eighth Amendment "had in common was the deliberate infliction of pain for the sake of pain-'superadd[ing]' pain to the death sentence through torture and the like" (alteration in original)).

194. *E.g.*, Vacco v. Quill, 521 U.S. 793, 807 (1997) (clarifying that the right to refuse medical treatment rests in the right "to bodily integrity and freedom from unwanted touching").

195. *E.g.*, Tennessee v. Lane, 541 U.S. 509, 534 (2004) (Souter, J., concurring) (describing the now discredited but "once-pervasive" forced sterilization of persons with disabilities).

On the other hand, mandatory vaccination does not entail a great degree of pain, the pain felt does not last very long, and the cultural meaning of being marked as a person with or without a vaccination is not very significant to most persons' identity. Thus, no fundamental right to refuse vaccination should stem from the special relationship of the body to rights. This does not mean that rights to refuse vaccination could not stem from other sources, such as religious freedoms.[196] But the argument that rights to bodily integrity or control over one's body should entail rights to refuse vaccination are too fetishistic about the body, subscribing to the fiction that our bodies could ever be "free" from all intrusion that implicates dignity or autonomy.

B. Rights to Change One's Body—From Abortion to Cross Dressing

Applying the principle to the question of rights to proactively modify and treat one's body would also likely result in relatively familiar negative rights against state criminalization of those modifications and treatments, such as abortion,[197] or treatments that are necessary to avoid pain.[198] A right to purchase many forms of medical treatment, or even "upgrades" to "healthy" or "normal" bodies, would also likely fall within the scope of rights we would want to protect in order to promote cultural velocity.

In this way, any "bodily" right protected under cultural velocity would be quite different from the American constitutional right to bodily integrity, which distinguishes between rights to resist intrusion into the body and rights to modify one's own body. The United States Supreme Court has recognized a fundamental right to refuse life sustaining medical treatment, but has not recognized a fundamental right to contract for assistance in committing suicide through the use of medication.[199] But the cultural velocity principle would not make this distinction because the right is not grounded merely in the "integrity" of the body, nor even merely in avoiding physical pain.

---

196. *E.g.*, *In re* Sherr v. Northport–East Northport Union Free Sch. Dist., 672 F. Supp. 81, 91 (E.D.N.Y. 1987) (finding that under First Amendment, religious exemption to New York's mandatory inoculation program for school children must be extended to all persons who sincerely hold religious beliefs, and not just persons who are bona fide members of a recognized religious organization).

197. *See* Roe v. Wade, 410 U.S. 113, 163 (1973).

198. Some Supreme Court justices have indicated that they would find constitutional fault with laws that prohibit the administration of pain medication necessary to alleviate great suffering, even in instances where the administration of the medicine is virtually certain to cause death. Washington v. Glucksberg, 521 U.S. 702, 736–37 (1997) (O'Connor, J., concurring) ("The parties and amici agree that in these States a patient who is suffering from a terminal illness and who is experiencing great pain has no legal barriers to obtaining medication, from qualified physicians, to alleviate that suffering, even to the point of causing unconsciousness and hastening death.").

199. *Compare* Cruzan v. Dir., Mo. Dep't of Health, 497 U.S. 261, 278 (1990), *with Vacco*, 521 U.S. at 807.

The principle also recognizes the importance of the body as a means of exploring and forming a cultural identity, and would therefore likely protect choices to engage in numerous kinds of body modifications, such as abortions, surgical and hormonal sex change procedures, cosmetic surgery, tattooing, piercing, and perhaps the ingestion of "mind altering" drugs. Again, rights are rarely absolute in any constitutional system, so professional and safety regulation of many of these practices would likely pass muster, as they do in the American constitutional system.[200]

The right to cross dress, and indeed, to dress as one pleases in public generally, would likely also be appropriately protected, even though dress and the use of jewelry do not involve the alteration of one's "natural" body. I have argued for such a freedom of dress in a prior work, but it is not on solid footing in American constitutional law.[201] Of course, as with all rights, sometimes the right will need to be balanced against other rights and other interests. Thus, threatening exercises of dress, such as indecent exposure committed with intent to harass and threaten, or the wearing of dangerous weapons as part of an "outfit," could be regulated, but the point remains that some kind of countermajoritarian right to manipulate one's appearance would make sense under the cultural velocity principle's application to the special importance of our bodies.

C. Rights to Use One's Body—from Sexual Liberty to Verbal Harassment

A right to sexual liberty[202] might also be grounded in this cultural velocity principle. Sexual activity is a use of the body that is largely described as having important components of physicality, just as the freedom to manipulate one's appearance through dress was historically described as a bodily freedom.[203] Moreover, sexual activity is one important means in many cultures of forming cultural affiliations and identity.

But would recognition of the body's importance to cultural velocity principle mean that we would protect every single use of the body? Could I claim that because my verbal harassment of another person involves use of my mouth that it is protected under the right? Could I claim that because thinking involves my brain, which is part of my body, an inquiry into my intent to commit a crime is an intrusion into my rights over my body?

---

200. For instance, the right to obtain an abortion may be limited by health and safety regulations throughout pregnancy. Planned Parenthood of Se. Pa. v. Casey, 505 U.S. 833, 874–78 (1992).

201. Ramachandran, *supra* note 9, at 16–17 (describing Supreme Court's denial of certiorari on the issue in the face of a circuit split).

202. *Cf.* Lawrence v. Texas, 539 U.S. 558, 567 (2003) (finding criminalization of same sex sodomy between consenting adults in private to violate the due process clause of the Fourteenth Amendment).

203. *See* James Allon Garland, *Breaking the Enigma Code: Why the Law has Failed to Recognize Sex as Expressive Conduct Under the First Amendment, and Why Sex Between Men Proves that it Should*, 12 L. & SEXUALITY 159, 180–87 (2003) (using the example of relationships between gay men to demonstrate the expressive nature of sex).

Recognition of the body's importance to cultural velocity, as a practical reason to select certain bodily-identity performative behaviors for legal protection, would not provide a reason for selecting speech or thought for legal protection. Those reasons would have to be articulated separately. This is because the cultural velocity conception has abandoned any idea that rights should map onto the physical borders of the human body. Thus, not every use or modification of the human body will necessarily implicate the body in a manner relevant to cultural velocity. Those uses that, as a cultural matter, tend to be experienced and understood as implicating subjectivity and cultural performance will be the uses that sit at the core of the right. This is why dress would be a more plausible candidate for protection than verbal speech, despite the fact that verbal speech involves a great deal of movement of one's mouth—part of the human body—while dress may not involve the movement of any part of the human body. Dress merely involves the placement of chattel upon the human body.

D. Patent Law

Beyond these mostly familiar negative rights against the state, applying the cultural velocity principle would also likely result in some rather unfamiliar forms of property regulation. For instance, it is possible that we might want to restrict end user license agreements ("EULAs") in certain forms of chattel because they would threaten individual innovation and ability to form and reform identity. We wouldn't ignore the implications of sellers retaining property rights in highly personal chattel just because it's chattel, rather than part of the organic body.

As one example, we might want to restrict the enforcement of EULAs against users of prostheses and other medical technology that intimately affects mobility, perception, or communication. This would include items like cochlear implants, robotic arms, visual aids, wheelchairs, and the like. While the idea that patent holders on these technologies would enforce EULAs against patients may seem far-fetched, the possibility that they would do so in a future in which many of these devices are "elective" is much stronger. If Apple is willing to enforce such agreements against purchasers of iPhones,[204] would it not consider enforcing them against purchasers of computer or phone implants? Memory card implants?

We might even want to simply restrict the enforcement of EULAs against users of any technology deeply tied to identity formation, whether or not it has anything to do with physicality or the body. It may be that

204. Katie Hafner, *Altered iPhones Freeze up*, N.Y. TIMES, Sept. 29, 2007 at C1, *available at* 2007 WLNR 19068410 ("Jennifer Bowcock, an Apple spokeswoman, said that when people went to update their software with their computer through iTunes, a warning appeared on the computer screen, making it clear that any unauthorized modifications to the iPhone software violated the agreement that people entered into when they bought the phone.").

the importance of the body will start to drop away, over time, and will no longer be a pragmatic, culturally appropriate way of singling out some identity formative practices from others for legal protection. The number of persons who find their phones, Facebook accounts, and other items to be more foundational to identity than even a prosthesis or their clothing may be growing. The cultural velocity approach is well-suited to deal with these sorts of changing conditions because it is an empirically falsifiable, consequentialist approach.

E. Employment and Housing Law

Another relatively unfamiliar area in which we might want to promote cultural velocity as it relates to bodily identity performance is in the regulation of employment and housing contracts. The U.S. economic system provides very minimal welfare rights, and indeed, provides no welfare entitlements to those who are considered able to work but do not.[205] Thus, the process of entering into employment and housing contracts for most adults in the U.S. is essential. With respect to the workplace, statutes like the Civil Rights Act of 1964[206] have made the American workplace more integrated and diverse along culturally salient categories such as race than many other private institutions such as churches or social clubs.[207] The workplace is already a venue where Americans can learn about those with different identities, where we can be challenged.[208]

We have already made statutory moves toward accommodating differences in bodies in both the employment and housing contexts, and we do not limit ourselves only to those differences in bodies which are fully biologically determined. For instance, we forbid discrimination on the basis of race, sex, and skin color[209] in these contexts, even though both race and sex can be thought of as socially constructed in part. We also require reasonable accommodation, at least in principle, of persons with disabilities at work and in school.[210]

---

205. The Personal Responsibility and Work Opportunity Reconciliation Act of 1996 permitted states to stop paying benefits to those otherwise eligible for welfare assistance if the recipient did not engage in qualified work. Pub. L. No. 104-193, § 407, 110 Stat. 2105, 2129–34 (codified as amended at 42 U.S.C. § 607 (2006)).

206. *See generally* 42 U.S.C. §§ 2000e-1 to -17 (2006).

207. *See* Cynthia L. Estlund, *Working Together: The Workplace, Civil Society, and the Law*, 89 GEO. L.J. 1, 17 (2000) ("[E]ven the partial demographic integration that does exist in the workplace yields far more *social* integration—actual interracial interaction and friendship—than any other domain of American society.").

208. *See* Cynthia L. Estlund, *Free Speech and Due Process in the Workplace*, 71 IND. L.J. 101, 112 (1995) ("The workplace functions not only as a self-governing institution and as a regulated institution; it also functions as a crucial intermediate institution that stands between the individual and the state.").

209. *See generally* 42 U.S.C. §§ 2000e, 3601–3619 (2006).

210. *See generally* 42 U.S.C. §§ 12101–12213 (2006).

However, we may want to expand this to accommodating not just biological or "accidental" difference, but also other differences in bodies. We may protect the ability of employees in their use of chattel closely related to the body, such as clothing, jewelry, or hairstyling.[211] We might also protect transgender employees, who change the gender presentation of their bodies through both surgical and somewhat more superficial changes. We could also protect transgender persons and persons who dress unusually from being excluded from housing. When it comes to this kind of private regulation of rights, a workable way of protecting rights in this private context will of course require a great deal of finesse. One appropriate way to balance the many interests at stake would be to require reasonable accommodation of employee differences in dress in most workplaces and provide statutory exceptions for certain jobs such as modeling or acting.[212]

Protections from sexual orientation discrimination could be conceived of as a means of promoting the right as well, as sexual conduct is certainly in the United States a means of exploring and forming identity through the use of the body.

Finally, we could consider expanding the kinds of bodies we reasonably accommodate in workplaces and schools beyond those currently protected by the Americans with Disabilities Act. Perhaps a body need not be "disabled" to the point of impairment in a "major life function" for us to recognize that it represents a different identity we may want others to interact with and learn from.[213]

Under this framework, persons who choose not to or cannot afford to "upgrade" their bodies through the use of drugs (such as steroids) and other biotechnology could be protected against unjust workplace or educational exclusion through reasonable accommodation of their comparative disadvantages and differences. However, the protections would not necessarily be great—"reasonable accommodation" is, after all, a relatively weak mandate. As a result, those who did not alter their bodies and therefore became less adept at truly core job functions would likely be left unprotected. I choose the weak requirement of reasonable accommodation purposefully as an example of what protecting rights to have a different body in a private sphere such as the workplace might look like.

---

211. I have proposed just this in a prior article, *Freedom of Dress*. *See* Ramachandran, *supra* note 9, at 37–43.

212. *Id.* at 61–64.

213. The Supreme Court has interpreted the Americans with Disabilities Act ("ADA") such that the protected class of disabled persons it protects is quite narrowly defined. *E.g.*, Sutton v. United Air Lines, Inc., 527 U.S. 471, 488–89 (1999) (finding that severely myopic job applicants were not disabled within meaning of the Act because they could use corrective lenses, and therefore received no protection from discrimination by the defendant employer on the basis of the myopia). The ADA was recently amended, however, to clarify that the coverage should be broad. ADA Amendments Act of 2008, Pub. L. No. 110-325, § 4, 122 Stat. 3553, 3555–56 (codified as amended at 42 U.S.C.A. § 12102 (West 2009)).

This relatively weak protection takes seriously the fact that some degree of coercion over bodies is inevitable. There will likely be economic incentives to take enhancing medication or alter one's body in other ways in the future, just as there are economic incentives to obtain cosmetic surgery that makes one more attractive now, and economic incentives to wear contact lenses or even obtain LASIK surgery. Because this coercion is inevitable, the aim of property and contract regulation to protect rights in the body can never be to provide absolute, substantive equality or absolute "freedom" for different bodies. The aim can only be to soften the effects of capital and other sources of private power on the variety of bodies and identities that become part of our culture—to prevent the consolidation of power over culture through power over bodies.

F. Welfare Law

With this in mind, we may be faced with body modifications in the future that are so economically valuable we simply cannot accommodate those who do not obtain them. Such accommodation may not be as easy as accommodation of visual or hearing impairments in many professions. For instance, we might be faced with the development of technology that dramatically expands memory or other components of intellectual capacity.

Thus, we might find it simpler at times to just redistribute wealth more dramatically, as a means of ensuring that those without capital retain a degree of material security no matter what they do or do not do to their bodies. This is certainly not strong protection, since all it would likely guarantee is that having an unusual body will not lead to starvation or homelessness. But we might argue that more generous welfare benefits are required for precisely this reason.

It is also worth pointing out that currently, we do not even have protection against starvation and homelessness for persons with unusual bodies, persons who are challenging our social norms. Transgender persons, for instance, are so marginalized under American law that they often become homeless and their very existence essentially criminalized.[214] No welfare system ensures basic material security for transgender persons, as even sex segregated homeless shelters are often unable to provide security.[215]

As another reform of welfare law, we might also provide welfare entitlements to obtain some economically valuable body modifications,

214. Spade, *supra* note 123, at 751–52.

215. *Id.* at 752–53 (describing how a host of laws, including gender documentation requirements even for driver's licenses, create a lack of access to employment and housing, and unsafe conditions at homeless shelters for many transgender persons).

to ensure that those with fewer financial resources and their progeny are not locked out of future opportunity and cultural participation.[216]

In fact, some prominent members of the "transhumanist" movement have predicted that once extremely valuable technology modifying the body is available, the masses will simply "demand" a dramatic wealth redistribution.[217] This is meant to respond to egalitarian concerns that body modifying or improving technology could widen the gap between poor and rich too far.[218]

While I find this prediction to be overly optimistic, the core point is a fair one—perhaps generous welfare law is a more direct instrument with which to deal with the inequality that social and cultural coercion over bodies inevitably causes. This makes sense once we have recognized that private property rights already threaten persons' control over their bodies. If that is the case, welfare law, rather than rights to "bodily integrity" or "personhood," may sometimes be the most effective way of protecting the ability to have a different or unusual body.

G. Body Commodification—From Prostitution to Organ Selling

Are there some sales of the body's parts or uses that must be prohibited, or that cannot be prohibited? There are some sales of the body or its parts that we might legitimately be concerned with from the perspective of protecting the public interest in cultural velocity. This is because we might legitimately be concerned that those with economic power and other forms of private influence could obtain too much control over culture by obtaining too much control over bodies. On the other hand, not every commodification of the "natural" body need be categorically forbidden, because not every commodification of the natural body impli-

---

216. The United Nations Convention on The Rights of Persons with Disabilities represents an important move towards welfare rights to certain aids. Section (g) of Article 4 states an obligation of parties to "undertake or promote research and development of, and to promote the availability and use of new technologies, including information and communications technologies, mobility aids, devices and assistive technologies, suitable for persons with disabilities, giving priority to technologies at an affordable cost." G.A. Res. 61/106, art. IV, U.N. Doc. A/RES/61/106 (Jan. 24, 2007). Article 20 states that parties "shall take effective measures to ensure personal mobility with the greatest possible independence for persons with disabilities." *Id.* at art. XX. This includes "[f]acilitating access by persons with disabilities to quality mobility aids, devices, assistive technologies and forms of live assistance and intermediaries, including by making them available at affordable cost." *Id.*

217. HUGHES, *supra* note 45, at 214–15.

218. *Id.*:

> It is unlikely that a future majority of service-providing 'commoners' with more free time, communications and democracy than today would tolerate being lorded over by a dynasty of non-working hereditary capitalists. They would vote to change the system. The trend in the social democracies has been to equalize income by raising the standards of the poorest as high as the economy can bear. In the age of robots, that minimum will be very high.

(internal quotation marks omitted) (quoting HANS MORAVEC, ROBOT: MERE MACHINE TO TRANSCENDENT MIND 132 (2000)).

cates this concern. But are there any sales of body parts or uses that must be *permitted*, from the perspective of protecting cultural velocity?

1. Prostitution

The best example of body commodification that arguably must be permitted under a proper understanding of the relationship of the body to rights is prostitution. Prostitution does not inherently involve pain, nor does it inherently involve indentured servitude or a use of the body that is extremely long in duration. Moreover, the long term physical effects of sexual activity, when engaged in with protections such as condoms, may be less significant than the long term physical effects of other forms of labor we permit, such as repetitive factory work.

Of course, the cultural and social meaning of engaging in sexual activity with another can be very significant, but forbidding all profits from activities that involve identity exploration, or from culturally meaningful acts, is unlikely to "protect" these activities in an otherwise essentially capitalist context. It is more likely to simply disincentivize them. This is why the United States Supreme Court has recognized that forbidding profits altogether from certain kinds of speech is usually an overbroad restriction on that speech.[219] It is also why, when I argued for a freedom of dress, I did not argue that modeling and acting should be outlawed.[220] Thus, it is arguable that not only should we abandon the claim that prostitution violates fundamental rights to bodily integrity, but we should also recognize a fundamental right to engage in prostitution, given the importance of sexual activity to identity formation and culture.

One might ask, does sex really need to be incentivized, the way certain forms of speech need to be incentivized? Do we really need financial incentives for sexual identity formation and exploration? The kind of sex that is culturally "normal" probably needs little additional financial incentive for persons to experiment with it. But some kinds of unusual or culturally challenging sexual experiences will probably happen more often in the United States if prostitution is legalized. For instance, persons with certain disabilities and disfigurements may be more likely to have sexual experiences if prostitution is legalized.[221] Alternately, some of these culturally challenging practices might be made more culturally visible, and actually have some effect on cultural norms regarding sex-

219. United States v. Nat'l Treasury Employees Union, 513 U.S. 454, 468–70 (1995).

220. Indeed, I proposed categorically exempting these jobs from reasonable accommodation of dress requirements. Ramachandran, *supra* note 9, at 63–64.

221. *See* Lucas, *supra* note 22, at 253 ("[I]n our nonideal world some individuals may be limited to a choice between commodified sex and involuntary celibacy."). *See generally* Elizabeth F. Emens, *Intimate Discrimination: The State's Role in the Accidents of Sex and Love*, 122 HARV. L. REV. 1307, 1385 (2009) (including information on state subsidies in some countries for the purchase of prostitution services by persons with disabilities).

uality, if prostitution is legal and more persons can "admit" to having sexual experiences with persons that are culturally disfavored sexually.[222]

Of course, sexual activity can be experienced as a bodily intrusion that affects one's very identity, even one's capacity to develop a different identity in the future.[223] Thus, many forms of regulation of prostitution may be wholly appropriate, and because any right to prostitution would not be grounded in "autonomy," such regulation would not necessarily infringe upon the right. Some examples of appropriate regulation could be laws that seek strong assurance of consent, laws requiring safe practices to prevent disease and pregnancy, laws ensuring the ability of parties to change their minds, and perhaps periodic surveillance or inspection to prevent violence. Even taxing the proceeds of prostitution and using the funds for mandatory counseling or other activities designed to ensure that the identity of the prostitute is not destroyed might be appropriate.

But in the face of arguments from sex workers that the work need not be destructive to identity formation and reformation, and in the face of the fact that we permit adults to take these risks with their physical bodies in the context of unpaid sex, prostitution appears to be a commodification of the body that we ought to legalize under a practical application of the cultural velocity principle, even one that accounts for the importance of sexual practice as a use of the body to identity formation. Arguably, some legal form of prostitution is even required as a matter of fundamental rights.

2. Organ Selling

Organ selling, in contrast, is a form of body commodification that involves lasting physical effects on the body. Thus, societies may be rightly concerned that those who sell organs might be selling away an aspect of their very subjectivity, in a manner that could affect their ability to explore many possible identities and cultural affiliations. But this is likely not the case for every organ or organ sale.

Selling one's eyes, a hand, or another such body part intimately involved in sensory perception or mobility may have lasting effects on one's identity that the seller regrets. On the other hand, selling a kidney seems less problematic with respect to this public interest in cultural evolution and change. Certainly the loss of a kidney entails future medical risk and even perhaps some experiences of pain that last well into the future. Thus, there are reasons to regulate such sales heavily to ensure

222. *Cf.* Lucas, *supra* note 22, at 261 ("Another benefit of commodification would be that . . . prostitutes could openly discuss their work, describe its pros and cons . . . . Customers could frankly discuss their experiences, both in the crass, economic way that commodification critics fear but also in honest terms about . . . what needs they seek to fill.").

223. *See id.* at 257–58.

fully informed consent. But a categorical prohibition on kidney sales does not appear to further the legitimate public interest in protecting individuals from contracting away their ability to form diverse identities.

On the other hand, there doesn't appear to be any reason to categorically *protect* a right to sell organs under this approach, either. Organ donation is not a particularly culturally salient practice, nor does it seem to involve physical experiences that are part of identity exploration, the way sexual acts arguably do. Moreover, most people donating organs are anesthetized during the process—it is not a bodily experience in which the donor is generally attempting to change or control his or her subjectivity, even temporarily. Thus, it is neither particularly "personal" in the sense of importance to subjective experience, nor particularly "political" in the sense of identity performative.

There may be traditional utilitarian concerns that would lead to bans on organ selling and other commodifications of the body, such as selling blood. The economic stability of a community and public health concerns are two potential reasons. And there may be traditional utilitarian concerns that would lead to *permitting* organ selling or other commodifications such as selling blood—we might want to increase the supply of organs, for instance. But these concerns need not be couched in terms of a fundamental right to bodily integrity, human dignity, or autonomy.

H. Animal-Human and Plant-Human Chimeras

What about the patenting of animal-human or plant-human chimeras, or transgenic entities,[224] that has raised so much new concern over assigning the legal status of property to persons or their parts?[225]

If an animal-human or plant-human chimera is capable of cultural interaction and affiliation, or of identity group formation and reformation, then it would seem that the cultural velocity principle would require us to forbid patent owners from controlling these entities in these activities.

It is important to note here that because of the way I have grounded the rights we protect under law, I am not requiring these entities to have "human intelligence" or to be capable of full political participation in order to be possessed of rights. Thus, my theory also protects almost all persons with disabilities, including those with mental disabilities. Even if a person with a severe mental disability never develops to the stage of traditional political participation, he is capable of identity performance and cultural affiliation, as well as of social attachments that are instruc-

224. *See* Munzer, *supra* note 57, at 124 (providing definitions of these terms).

225. *See, e.g.*, Bagley, *supra* note 2, at 546; *see also supra* notes 50, 60, 62–64, 66 and accompanying text.

tive and useful to society.[226] But some minimal level of "intelligence" or cognitive processing is probably required for a being to attain the kinds of identities or bodily identity performances that could contribute to our cultural evolution.

Although a proper conception of how the body relates to rights may ultimately prevent a patent-holder from exercising control over the beings she creates, the principle I have articulated would not necessarily prevent a patent holder from preventing another person or company from producing identical beings for the duration of the patent. This portion of the patent right does not seem to lastingly affect our interest in cultural evolution or cultural velocity. Thus, a patent holder in an animal-human chimera or transgenic creature might see his or her patent rights diminished, but not destroyed, under a proper conception of the body's relationship to fundamental rights.

It may seem strange for a company to hold a patent in the method of producing an entity that has intelligence, interacts with others, and is a member of society. But this state of affairs is not without analogy in laws with which we are familiar. Not long ago, children were treated like the property of their fathers.[227] But parents do not exercise absolute dominion over their children. For instance, parents cannot abuse or neglect their children in the United States. And even in the past, fathers had duties of care to their children.[228] Moreover, once children reach adulthood, parents exercise no dominion or control over their children at all. And yet, this recognition that parents should not have total control over their children need not stop us from using property rights to govern control over embryos that parents have contributed to genetically. Allowing a parent to sell an embryo or donate rights in the embryo to another parent is not the same as allowing a parent to exercise dominion over a child.[229]

CONCLUSION

The problem of property and freedom of contract threatening fundamental rights is not new. However, the physical merger of humans and non-humans is about to become prominent. Thus, we can no longer subscribe to the fiction that by drawing a sharp line between humans and the

226. *See* NUSSBAUM, *supra* note 79, at 99, 129 (describing her "conception of the person as a social animal, whose dignity does not derive from an idealized rationality," and also "the advantage of understanding humanity and its diversity that comes from associating with mentally disabled people on terms of mutual respect and reciprocity").

227. *See* MARY ANN MASON, FROM FATHER'S PROPERTY TO CHILDREN'S RIGHTS: THE HISTORY OF CHILD CUSTODY IN THE UNITED STATES 1–13 (1994) (describing the fathers' custody over children as similar to the master/servant relationship, which involved duties of care, as well as the ability to contract out and profit from a servant's labor).

228. *Id.* at 12–13.

229. *See* Dan L. Burk, *Patenting Transgenic Human Embryos: A Nonuse Cost Perspective*, 30 HOUS. L. REV. 1597, 1648 (1993) ("[T]he holder of a patent for a transgenic human being could presumably prevent others from making, using, or selling such a transgenic human being, but this does not mean that the patent holder could impress the patented person into servitude or bondage.").

rest of the world, and prohibiting the commodification of the human, we have sufficiently protected the need for diverse exercises of human agency against property rights. Property rights and freedom of contract have always threatened to stifle us via our bodies, such as by constructing us as "disabled" rather than "abled," or by determining whether we will need to perform physical labor to survive.

We must accept that the line between humans and property is being blurred, and that already human self-control can be affected by the distribution of property rights. To deal with threats to physical self-control we cannot rely on the formal sanctity of the human body as separate from property. At the same time, we cannot rely solely on unregulated, autonomy-based, property-like rights in the body. For the same reasons that all forms of property are regulated, and do not entail sole dominion, property-like rights in one's body should also sometimes be regulated.

We instead need a conception of rights that would simultaneously protect the diverse set of bodily choices that form individuals' identities and cultural affiliations, while using the regulation of property and contract to protect against a few actors obtaining monopolies over the subjectivities and identities of others. But because we are at least in part products of technology, our rights to control our own subjectivities or to form diverse identities cannot be protected by any kind of neat "right to bodily integrity" or autonomy over the natural, human body.

RETHINKING "INSURANCE," ESPECIALLY AFTER AIG

BOBBY L. DEXTER†

INTRODUCTION

In order to qualify as an "insurance company" for federal income tax purposes, both life insurance companies and their property and casualty counterparts must clear the hurdle of Section 816(a) of the Internal Revenue Code ("the Code").[1] That section clarifies that insurance company status follows only if more than half of the entity's business is the issuance of insurance or annuity contracts.[2] Although neither the Code nor the Treasury Regulations define "insurance,"[3] longstanding common law doctrine dictates that an arrangement will constitute insurance only if it incorporates requisite risk shifting and risk distribution.[4] A life insurance policyholder, for example, typically shifts the financial risk of his untimely demise to an insurance company by paying premiums.[5] In turn, the life insurance company broadly distributes that risk by collecting premiums from a large population of policyholders[6] so that any claims presented will not present a financial challenge to the company relative to the aggregate premiums received and set aside for claim payments.[7] In

---

† Associate Professor of Law, Chapman University School of Law. B.A., Yale University, 1989; J.D., Harvard Law School, 1992. I would like to thank Chapman University School of Law for providing research support for this article and my faculty colleagues for offering comments and suggestions during the presentation of this work. Thanks are also in order for comments from participants attending the Southern California Junior Law Faculty Workshop held at Pepperdine School of Law, the 2009 Annual Meeting of the Law & Society Association, and the 2009 Junior Tax Scholars Workshop held at Brooklyn Law School.

1. *See* I.R.C. § 816(a) (2006). With respect to property and casualty companies subject to tax under § 831, section 831(c) indicates that "insurance company" under § 831 has the same meaning as that set forth in § 816(a). I.R.C. § 831(c) (2006).

2. *See* § 816(a). A company may also qualify by reinsuring risks underwritten by other insurance companies. *See id.*

3. *See* Rev. Rul. 2005-40, 2005-2 C.B. 4, 2005 WL 1415557.

4. *See* Helvering v. Le Gierse, 312 U.S. 531, 539 (1941) ("Historically and commonly insurance involves risk-shifting and risk-distributing.").

5. *See* Clougherty Packing Co. v. Comm'r, 811 F.2d 1297, 1300 (9th Cir. 1987) ("Shifting risk entails the transfer of the impact of a potential loss from the insured to the insurer. If the insured has shifted its risk to the insurer, then a loss by or a claim against the insured does not affect it because the loss is offset by the proceeds of an insurance payment."); Spring Canyon Coal Co. v. Comm'r, 43 F.2d 78, 80 (10th Cir. 1930) (noting that the taxpayer would have shifted the risk had it paid the premium).

6. *See Clougherty Packing Co.*, 811 F.2d at 1300 ("Insuring many independent risks in return for numerous premiums serves to distribute risk.").

7. *See id.* As an interesting historical aside, several of the initial founders of the North Carolina Mutual Life Insurance Company (which currently has insurance in force exceeding $12 billion and is the oldest and largest historically African-American-owned life insurance company in the United States, *see* N.C. Mut. Life Ins. Co., http://www.ncmutuallife.com/newsite/pages/about.html (last visited Sept. 7, 2009)) had to pool their own funds after a quick meeting in the rear of a barber shop in order to pay the $40 death benefit claim of a widow. *See* 2 ENCYCLOPEDIA OF AFRICAN AMERICAN BUSINESS 519 (Jessie Carney Smith ed., 2006).

adopting that approach, the company takes advantage of the so-called "law of large numbers." Using reliable probability data, it can generally coordinate premium receipts with predicted losses.[8] With a population of policyholders in a life insurance pool sufficiently large and varied, *predicted* mortality rates will roughly correlate with the population's *actual* mortality rates,[9] and the company can generally pay benefits as they come due over the long term without going broke in the short term.[10]

For many companies, satisfying the qualitative and quantitative "insurance" standards presents no problem. Accordingly, the company does not question its status as an insurance company. Logically, those policyholders entitled to deduct premiums paid to insurance companies as ordinary and necessary business expenses also have no reason to question the tax deductibility of the payments made.[11] Historically, however, there are prominent victims of the risk shifting/distributing standard, entities known in the business world as "captive" insurance companies.[12] Generally speaking, a captive insurance company (which is routinely organized and operated as a wholly-owned subsidiary of a parent entity) is often created to provide coverage to a business that cannot secure the insurance it needs from the marketplace, or can only secure such insurance at a prohibitive cost[13] (for example, a company operating nuclear reactors in an urban area and needing insurance against the risk of a meltdown). Legitimate business needs notwithstanding, the Internal Revenue Service ("IRS" or "Service") has long been hostile to the notion that a wholly-owned subsidiary can insure its parent company[14] and, accordingly, that the payment of a premium from a parent to its captive insurer subsidiary qualifies as a deductible ordinary and necessary business expense. Attempts to resolve this issue have, not surprisingly, resulted in substantial judicial activity, a series of IRS pronouncements, and a healthy body of professional and scholarly commentary.[15] For some time, the Service

8. *See Clougherty Packing Co.*, 811 F.2d at 1300.

9. *See id.* (linking the law of large numbers, a statistical phenomenon, and the concept of risk distribution).

10. *See id.* ("Distributing risk allows the insurer to reduce the possibility that a single costly claim will exceed the amount taken in as a premium and set aside for the payment of such a claim. . . . By assuming numerous relatively small, independent risks that occur randomly over time, the insurer smoothes out losses to match more closely its receipt of premiums.").

11. *See* Treas. Reg. § 1.162-1(a) (as amended in 1993) ("Among the items included in business expenses are . . . insurance premiums against fire, storm, theft, accident, or other similar losses in the case of a business . . . .").

12. For a basic definition of "captive insurance company," see Gordon A. Schaller & Scott A Harshman, *Use of Captive Insurance Companies in Estate Planning*, 33 AM. C. TR. & EST. COUNS. J. 252, 252 (2008). Application of the traditional "insurance" test to captive insurance companies is discussed in Part II of this Article.

13. *See id.*

14. *See, e.g.*, Rev. Rul. 77-316, 1977-2 C.B. 53, 1977 WL 43573, *obsoleted by* Rev. Rul. 2001-31, 2001-1 C.B. 1348, 2001 WL 606232. As discussed in greater detail in Part II, the prior ruling concluded, inter alia, that a parent had not entered into an insurance relationship with its wholly-owned insurance subsidiary.

15. *See, e.g.*, Malone & Hyde, Inc. v. Comm'r, 62 F.3d 835, 840–43 (6th Cir. 1995); Ocean Drilling & Exploration Co. v. United States, 988 F.2d 1135, 1153–54 (Fed. Cir. 1993); Amerco, Inc.

reasoned that a premium payment from one related entity to another in its affiliated or consolidated group for insurance simply did not shift any risk because, in the end, the financial burden of loss coverage did not leave "the economic family."[16] Without risk shifting, no insurance arrangement existed, and a tax deduction for a premium payment from the entity to its fellow group member was deemed inappropriate.[17] Even in the wake of judicial attack[18] and scholarly criticism,[19] the Service persistently appeals to traditional risk shifting and risk distributing standards with regard to the tax deductibility of insurance premium payments.[20]

Whether stating its opposition to a given arrangement under the rubric of risk shifting/distributing or "economic family," the Service seeks, at root, to prevent the taking of a deduction for what it sees as merely a contribution to a contingency reserve. Its opposition is not unwavering

---

v. Comm'r, 979 F.2d 162, 164, 168 (9th Cir. 1992); Harper Group v. Comm'r, 979 F.2d 1341, 1342 (9th Cir. 1992); Sears, Roebuck & Co. v. Comm'r, 972 F.2d 858, 860–62 (7th Cir. 1992); Humana, Inc. v. Comm'r, 881 F.2d 247, 251–52 (6th Cir. 1989); Clougherty Packing Co. v. Comm'r, 811 F.2d 1297, 1305–07 (9th Cir. 1987); Beech Aircraft Corp. v. United States, 797 F.2d 920, 922–23 (10th Cir. 1986); Carnation Co. v. Comm'r, 640 F.2d 1010, 1013 (9th Cir. 1981); Hosp. Corp. of Am. & Subsidiaries v. Comm'r, 74 T.C.M. (CCH) 1020, 1997 WL 663283, at *30, *35 (1997); Kidde Indus., Inc. v. United States, 40 Fed. Cl. 42, 56–58 (1997), *appeal dismissed per stipulation*, 194 F.3d 1330 (Fed. Cir. 1999) (unpublished table decision); Gulf Oil Corp. v. Comm'r, 89 T.C. 1010, 1025–29 (1987); Rev. Rul. 2008-08, 2008-1 C.B. 340, 2008 WL 131946; Rev. Rul. 2005-40, 2005-2 C.B. 4, 2005 WL 1415557; Rev. Rul. 2001-31, 2001-1 C.B. 1348, 2001 WL 606232; Rev. Rul. 78-338, 1978-2 C.B. 107, 1978 WL 41909; Rev. Rul. 77-316, 1977-2 C.B. 53, 1977 WL 43573, *obsoleted by* Rev. Rul. 2001-31, 2001-1 C.B. 1348, 2001 WL 606232; I.R.S. Priv. Ltr. Rul. 96-24-028 (June 14, 1996), 1996 WL 326435; James A. Christopherson, *The Captive Medical Malpractice Insurance Company Alternative*, 5 ANNALS HEALTH L. 121, 125–32 (1996); Karen Gantt, *Federal Tax Treatment of Medical Malpractice Insurance Alternatives for Nonprofits*, 52 DRAKE L. REV. 495, 514–21 (2004); Ray A. Knight & Lee G. Knight, *Disregarding the Separate Corporate Entity of Captive Insurance Companies: A Violation of the* Moline Properties *Doctrine*, 14 J. CORP. L. 399, 405–09 (1989); Robert W. Minto, Jr., *Captives and RRG'S in the Reinsurance Environment*, 889 PLI/COMM 837, 843–44 (2006); Schaller & Harshman, *supra* note 12, at 254–58; F. Roy Sedore, *Insurance Premium Deductibility*, 779 PLI/TAX 1061, 1067–69 (2007); Stuart R. Singer, *When the Internal Revenue Service Abuses the System: Captive Insurance Companies and the Delusion of the Economic Family*, 10 VA. TAX REV. 113, 160–63 (1990); Scott A. Taylor, *Taxing Captive Insurance: A New Solution for an Old Problem*, 42 TAX LAW. 859 *passim* (1989); Donald Arthur Winslow, *Tax Avoidance and the Definition of Insurance: The Continuing Examination of Captive Insurance Companies*, 40 CASE W. RES. L. REV. 79, 110–14 (1990); Armando Gomez, Note, *A Practical Approach to the Captive Insurance Problem*: Sears, Roebuck & Co. v. Commissioner, 46 TAX LAW. 619, 620–23 (1993).

16. *See* Rev. Rul. 77-316, 1977-2 C.B. 53, 1977 WL 43573, *obsoleted by* Rev. Rul. 2001-31, 2001-1 C.B. 1348, 2001 WL 606232.

17. *See id.*

18. *See, e.g.*, Sears, Roebuck & Co. v. Comm'r, 972 F.2d 858, 861 (7th Cir.) (indicating that in *Le Gierse* the Supreme Court "was not writing a definition for all seasons and had no reason to"). The court went on to say that "it is impossible to see how risk shifting can be a *sine qua non* of 'insurance.'" *Id.* at 862.

19. *See, e.g.*, Singer, *supra* note 15, at 119 (describing the risk shifting/distributing standard as dicta from dated U.S. Supreme Court authority and the Service's economic family theory as an invention); Knight & Knight, *supra* note 15, at 417 (arguing that the "economic family" theory—advanced by the Service in concluding that risk shifting and distributing were absent in some transactions—cannot be reconciled with the *Moline Properties* doctrine); Note, *The New York Stock Exchange Gratuity Fund: Insurance That Isn't Insurance*, 59 YALE L.J. 780, 782 (1950) (describing the risk shifting/distributing test as "cryptic").

20. *See, e.g.*, Rev. Rul. 2008-08, 2008-1 C.B. 340, 2008 WL 131946; Rev. Rul. 2005-40, 2005-2 C.B. 4, 2005 WL 1415557.

because various Code provisions explicitly sanction the taking of deductions for additions to reserves for *some* companies. For example, life insurance companies[21] facing ultimate liability for a given death benefit and nuclear energy companies[22] facing decommissioning obligations may properly deduct premium payments made into their own reserves. Yet, despite the fact that establishing substantial financial reserves is a wise move for any responsible company, the ability to do so at the expense of the United States Treasury is not a matter of easily obtained Congressional grace. Should it be?

This Article asks and attempts to answer questions of doctrine and policy. Should the "reserving function" be limited, as a matter of sound tax policy, to a discrete subset of taxpayers operating in industries of relative certainty (e.g., death, decommissioning, and the like), or should Congress allow deductions for limited contributions to contingency reserves more broadly? If one deems a general contingency reserve regime fiscally untenable, then can one rationally consider allowing gradual, limited, and regulated reserving for companies "too big to fail?" Or, does that approach ask way too much for financial exigencies that merely *might* occur? Also, as a purely doctrinal matter, should the Service altogether abandon its traditional, yet utterly shakable, risk shifting/distributing approach, or reserve that approach for companies that could be served well enough by the established insurance markets but merely wish to use a captive insurance company to reduce overall operating expenses (i.e., leave *this* to experts, and don't try it at *your* office)?

Given the posture of various courts and commentators, one can certainly argue the case for tweaking the "insurance" test to accommodate individual entities (or affiliated groups) with risks that are both statistically "certain-to-occur-at-some-point" and sufficiently numerous, homogenous, and independent[23] to take advantage of the law of large numbers without the involvement of an entity outside the affiliated group (for example, a single corporate entity or group with a massive fleet of service vehicles needing some form of insurance coverage).

In light of recent economic challenges,[24] now may be a good time to reconsider the permissible ambit and optimal scope of the reserving function in the tax arena. We live in a time when one hungry corporate Goliath grumbles impatiently behind its household name big brother, each

21. I.R.C. § 805(a)(2) (2006).

22. *See* I.R.C. § 468(a)(1)(A) (2006).

23. *See generally* Gulf Oil Corp. v. Comm'r, 89 T.C. 1010, 1025 n.9 (1987) ("As a theoretical matter, risk distribution or pooling requires: (i) Mass - * * * , (ii) Homogeneity - * * *, and (iii) Independence - * * *. If these requirements are met to some minimum extent, the principles of average and large numbers operate . . . .").

24. *See* Adam Zagorin & Michael Weisskopf, *Inside the Breakdown at the SEC*, TIME, Mar. 9, 2009, at 34, *available at* 2009 WLNR 3721270 (describing current economic conditions as "the worst financial crisis since the Great Depression").

queuing for its slice of federal bailout.[25] Neither is alone. Americans are also grumbling. They rightfully ask why they should be on the hook for resuscitating companies that should have never been allowed to wade so far into the deep.[26] With justified skepticism, they question the wisdom of bailing out companies with their hard-earned tax dollars[27] while corporate executives, without the restraining leash of rapt oversight,[28] recline on private jets and make their way to the next boondoggle hoping that the redecoration of their offices will have been completed by the time they get back.[29] Given the potential negative economic ramifications for the broader U.S. economy, maybe certain companies were indeed "too big to fail"[30]—but certainly they were not too big to strategically reserve some portion of their profits instead of routinely showering millions[31] on a select few for contract-based or "performance-based"[32] compensation. Would it not be better for American taxpayers to collectively suffer corporate deductions for the gradual creation of a *limited*

25. *See, e.g.*, Bill Saporito, *Is This Detroit's Last Winter?*, TIME, Dec. 15, 2008, at 34, *available at* 2008 WLNR 23306940 (indicating that American International Group, Inc. ("AIG") and Citigroup had received multi-billion dollar bailouts and that Ford, General Motors, and Chrysler were hoping to receive federal bailouts). Some companies have returned several times for additional assistance. *See* Francesco Guerrera, *AIG Considers Break-up in Bid to Stay Afloat*, FIN. TIMES, Feb. 25, 2009, at 1 (reporting that AIG might be broken up as part of a radical restructuring of the company and noting that, at the time, the government had bailed out AIG three times in five months).

26. *See, e.g.*, Saporito, *supra* note 25, at 34 (highlighting situational irony by pointing out that "AIG torpedoes the entire economy and gets a $150 billion handout; Citigroup takes risks no sane manufacturing company would even contemplate and is rewarded with a $20 billion federal bailout").

27. Current proposals direct trillions of dollars at recovery efforts. *See* John Fritze, *Trillions Aimed at Recovery*, USA TODAY, Feb. 11, 2009, at 1A, *available at* 2009 WLNR 2668766 ("The White House unveiled a sweeping proposal . . . to spend as much as $2 trillion in public and private funds to prop up the nation's financial system as the Senate narrowly approved an $838 billion stimulus intended to jump-start the failing economy.").

28. *See, e.g.*, Alan Beattie & Edward Luce, *Obama Gets Tough on Pay for Executives*, FIN. TIMES, Feb. 5, 2009, at 1 (noting U.S. President Barack Obama's disgust at various Wall Street excesses and reporting that Wall Street executives received $20 billion in bonuses in 2008 despite massive market losses); *see also* Sarah O'Connor, *Top Senators Rap Fed Over AIG Rescue*, FIN. TIMES, Mar. 6, 2009, at 3 (noting the warnings of the U.S. Senate's banking committee with respect to future financial support of AIG given the "failure to reveal how the $163bn injected into the insurance company has been spent").

29. *See* Mark DeCambre & Paul Tharp, *Goofy Getaway—AIG Ripped for Senseless Spending*, N.Y. POST, Oct. 10, 2008, at 34, *available at* 2008 WLNR 19335564; *see also* Terry Keenan, *The Reign of Thain Was Mainly Just a Pain*, N.Y. POST, Jan. 25, 2009, at 32, *available at* 2009 WLNR 1631465.

30. Many consider AIG the quintessential company that is "too big to fail." *See, e.g.*, Bill Saporito, *How AIG Became Too Big to Fail*, TIME, Mar. 30, 2009, at 24, *available at* 2009 WLNR 5204402.

31. *See, e.g.*, Mimi Hall et al., *Critics Blast AIG as Flap Escalates over Bonuses*, USA TODAY, Mar. 17, 2009, at 1A, *available at* 2009 WLNR 5033330 (reporting growing public disgust with corporate America and indicating that bonuses paid to certain AIG executives were negotiated in early 2008). Apparently, compensation restraint is now a reality with respect to certain federal funds recipients. *See* Beattie & Luce, *supra* note 28, at 1 (reporting that going forward, executives at companies receiving a certain level of federal assistance would receive no more than $500,000 per year as compensation).

32. Although I.R.C. § 162(m)(1) generally disallows a deduction in the publicly-held company context for "applicable employee remuneration" above $1,000,000, certain performance based compensation is excluded from the definition of "applicable employee remuneration" under § 162(m)(4)(C). I.R.C. § 162(m) (2006) (limiting deductions for compensation in some contexts).

contingency reserve with which the companies could head off full-scale financial havoc at the pass (or at least make a *hefty* contribution to their own bailout), rather than being forced to ante up a lump sum on short order, bear the brunt of the various agency costs inherent in the corporate form[33] (enhanced and exacerbated by gargantuan bailout infusions), and shoulder all of the attendant repayment risks at the same time?

Turning first to the doctrinal inquiry, Part I of this Article glances back to the common law origin of the risk shifting/distributing standard and reaches back a bit further, to highlight early authority in which the Service frowns harshly on taxpayer deductions for self-reserving activity.

To flesh out evolving doctrinal contours and set the stage for a critique, Part II shifts the focus to the captive insurance company context where much—if not most—of the litigation regarding risk shifting/distributing has taken place. In addition to addressing the rise and fall of the economic family argument empire, Part II also explains which captive insurance arrangements work and highlights the Service's most recent pronouncements that effectively delineate its current litigation position.

Part III briefly summarizes prior criticism of the risk shifting/distributing standards, both of which have been gradually weakened by courts and commentators, before going on to explore the question of whether recent IRS pronouncements reflect a new and aggressive elevation of form over substance in the captive insurance arena. Part III also notes that current business realities, coupled with the nefarious practices that many insurance companies habitually employ, compel some questioning of the utility of the risk shifting/distributing standards. Arguably, these standards best apply in more idealistic settings.

Part IV presents and assesses a two-pronged proposal for reform: the allowance of both contextual self-insuring (which seeks to modify the prevailing standard), and utilitarian contingency reserving (which would permit regulated contingency reserving using existing nuclear decom-

---

33. One commentator notes the following:

> Agency costs are defined as the sum of the monitoring and bonding costs, plus any residual loss, incurred to prevent shirking by agents. In turn, shirking is defined to include any action by a member of a production team that diverges from the interests of the team as a whole. As such, shirking includes not only culpable cheating, but also negligence, oversight, incapacity, and even honest mistakes.
>
> . . . .
>
> . . . [A]gency costs are the inevitable consequence of vesting discretion in someone other than the residual claimant.

STEPHEN M. BAINBRIDGE, CORPORATION LAW AND ECONOMICS 35–37 (2002). Directors and officers of companies are often protected from liability for mismanagement of the business and affairs of a given company. Indeed, those individuals often enjoy several layers of protection, given that they tend to be covered by both primary policies and several "excess" policies (i.e., those that will generally pay after primary policy limits have been reached). *See* Eric S. Connuck, *Excess D&O Insurance: The Exhaustion by Payment Condition*, BUS. L. TODAY, Sept.–Oct. 2008, at 45.

missioning reserve fund rules and procedures as a basic model). Part V sets forth concluding thoughts.

I. TRADITIONAL DEFINITIONS: "INSURANCE COMPANY" AND "INSURANCE"

Insurance companies, whether life and health companies or property and casualty companies, commonly have substantial amounts of both underwriting income (e.g., premium payments, fees, and assessments)[34] and investment income (e.g., interest, dividends, and rents).[35] At the same time, these companies, because they perform an important societal function, enjoy favorable treatment under the federal income tax laws. Among other things, they have the ability to take current deductions for increases in certain internal reserves.[36] Reasoning that an entity devoted primarily to the generation of investment income should not enjoy unduly favorable tax treatment, Congress requires that an entity seeking to qualify as an "insurance company" for federal income tax purposes satisfy both qualitative and quantitative standards. Speaking in general terms, the company must be in the business of issuing insurance or annuity contracts,[37] and that business activity must constitute more than half of all the company's total business.[38]

Although the "more than half" quantitative standard has long applied to life insurance company determinations, non-life companies (such as property and casualty) could, at least until recently, qualify as insurance companies by establishing that the issuing of insurance contracts was their "primary and predominant" business activity.[39] By way of the Pension Funding Equity Act of 2004,[40] Congress eliminated the quantitative standard duality and now requires that *all* insurance companies satisfy the "more than half" standard.[41] Unfortunately, Congress did not use that legislation as a vehicle to clarify the precise meaning of "insurance."

---

34. *See* KENNETH BLACK, JR. & HAROLD D. SKIPPER, JR., LIFE AND HEALTH INSURANCE 971 (13th ed. 2000).

35. *See* I.R.C. § 832(b)(2) (2006) (listing specific items of investment income).

36. Life insurance companies are entitled to a deduction for increasing their life insurance reserves. *See* I.R.C. §§ 805(a)(2), 807(b), (c)(1) (2006). Other insurance companies are allowed adjustments in calculating underwriting income for increases in unearned premiums and unpaid loss reserves. *See* § 832(b)(3), (b)(4)(B), (b)(5)(A)(ii).

37. The company may also qualify as an insurance company by reinsuring risks which are underwritten by insurance companies. I.R.C. § 816(a)(2) (2006). For a basic application of the qualitative standard, see Allied Fid. Corp. v. Comm'r, 572 F.2d 1190, 1194 (7th Cir. 1978) (holding that a wholly-owned subsidiary was not an insurance company because the character of its primary business, the writing of bail bonds, was not the writing of contracts of insurance).

38. § 816(a)(2).

39. Treas. Reg. § 1.801-3(a)(1) (as amended in 1972). Thus, it was simply easier for a property and casualty company to qualify as an insurance company. *See* H.R. REP. NO. 98-432, pt. 2, at 1402–04 (1984), *reprinted in* 1984 U.S.C.C.A.N. 697, 1047 (indicating Congress's intent to create a standard more strict and precise for life insurance company characterization).

40. Pension Funding Equity Act of 2004, Pub L. No. 108-218, § 206, 118 Stat. 596, 611 (codified as amended at I.R.C. § 831(c) (2006)).

41. *See* I.R.C. §§ 816(a)(2), 831(c).

Thus, as a traditional starting point, we are left with the risk shifting/distributing standards set forth in 1941, in the landmark case of *Helvering v. Le Gierse*.[42]

Less than a month before her death, 80-year-old Edyth Le Gierse purchased both a single-premium life insurance contract and an annuity from the Connecticut General Life Insurance Company ("Connecticut General").[43] Together, the contracts obligated Connecticut General to pay her $589.80/year for the remainder of her life and a $25,000 death benefit to her daughter, the designated beneficiary.[44] Edyth executed the contracts and paid $27,125 in a lump sum to Connecticut General.[45] Ultimately, it was hoped that these arrangements would allow the exclusion of the $25,000 death benefit from Edyth's gross estate for federal estate tax purposes.[46] Concluding that the death benefits should not have been excluded from Edyth's gross estate, the Commissioner asserted a deficiency, but was forced to pursue the case to the U.S. Supreme Court before achieving the desired result.[47]

Writing for the majority, Justice Murphy noted that the parties complied in all respects with standard contractual formalities,[48] but went on to emphasize that "the [purported insurance] amounts must be received as the result of a transaction which involved an actual 'insurance risk' at the time the transaction was executed. Historically and commonly insurance involves risk-shifting and risk-distributing."[49] Viewing the transaction holistically, the Court concluded that the life insurance contract would not have been issued without the annuity contract and that "[c]onsidered together, the contracts wholly fail to spell out any element of insurance risk."[50] The Court emphasized that all contract consideration was prepaid,[51] and any risk that the amounts paid by Edyth Le Gierse would not generate earnings sufficient to cover the annuity payments due her was an *investment* risk, not an insurance risk.[52] Accordingly, the Court held that the $25,000 should have been included in the decedent's gross estate.[53]

Although there is only so much to be learned about the fundamental nature of insurance from *Le Gierse*—namely, that insurance is not the

42. 312 U.S. 531 (1941).
43. *Id.* at 536.
44. *Id.*
45. *Id.*
46. *See id.* at 537–38.
47. *See id.* at 537.
48. *See id.*
49. *Id.* at 539.
50. *Id.* at 541.
51. *Id.* at 542.
52. *Id.*; *see also* Comm'r v. Treganowan, 183 F.2d 288, 290 (2d Cir. 1950) ("[T]here is no risk unless there is uncertainty or, to use a better term, fortuitousness.").
53. *See Le Gierse*, 312 U.S. at 542.

equivalent of a mere investment risk[54]—the opinion serves as landmark decision because it articulates the risk-shifting and risk-distributing standards.[55] Even though the *Le Gierse* opinion was criticized as cryptic soon after issuance of the opinion[56] and weakened, to some extent, by subsequent judicial developments[57] and scholarly commentary,[58] the two-part test has managed to survive for decades.[59] Whether this sturdy, conceptual immunity is a good thing, whether its application remains justifiable, and whether the standard has weathered steady criticism in worse condition than its proponents are willing to acknowledge, all remain open questions to be addressed in the following Parts of this Article.

There remains a need to separate true "insurance" companies entitled to favorable tax treatment from non-insurance companies that fail to qualify for favorable treatment. Thus, some gate-keeping standard must be in place. That standard must, however, be rational, and ideally should be fully consistent with sound tax policy. Further, compliance standards should be readily ascertainable, given that a company that unwittingly fails the standard may find itself in an unexpected or unworkable tax environment, and those paying purported premiums may suddenly find that deduction of the amount was improper.

Whether deservedly or not, captive insurance companies and their premium-paying organizers have historically been the predominant vic-

---

54. *See id.* For a more recent statement of this principle, see Rev. Rul. 89-96, 1989-2 C.B. 114, 1989 WL 550753 (emphasizing that the mere assumption of an investment risk cannot create an insurance contract for federal income tax purposes).

55. *See, e.g.*, Rev. Rul. 2008-08, 2008-1 C.B. 340, 2008 WL 131946 (citing *Le Gierse* and noting that it is United States Supreme Court precedent).

56. *See* Note, *supra* note 19, at 782. The Note also argues that in *Estate of Strauss*, the Tax Court erred in concluding that risk-shifting had not occurred and that the death benefit plan at issue did not constitute "insurance." *See id.* at 784–86. The Tax Court was later reversed by the Court of Appeals for the Second Circuit. *Treganowan*, 183 F.2d at 293.

57. *See* Malone & Hyde, Inc. v. Comm'r, 62 F.3d 835, 838–39 (6th Cir. 1995); Ocean Drilling & Exploration Co. v. United States, 988 F.2d 1135, 1144–46 (Fed. Cir. 1993); Amerco, Inc. v. Comm'r, 979 F.2d 162, 164–65 (9th Cir. 1992); Harper Group v. Comm'r, 979 F.2d 1341, 1342 (9th Cir. 1992); Sears, Roebuck & Co. v. Comm'r, 972 F.2d 858, 861–62 (7th Cir. 1992); Humana, Inc. v. Comm'r, 881 F.2d 247, 255 (6th Cir. 1989); Beech Aircraft Corp. v. United States, 797 F.2d 920, 922 (10th Cir. 1986); Carnation Co. v. Comm'r, 640 F.2d 1010, 1013 (9th Cir. 1981); Kidde Indus., Inc. v. United States, 40 Fed. Cl. 42, 56–57 (1997), *appeal dismissed per stipulation*, 194 F.3d 1330 (Fed. Cir. 1999) (unpublished table decision); Gulf Oil Corp. v. Comm'r, 89 T.C. 1010, 1025–27 (1987); Rev. Rul. 2008-08, 2008-1 C.B. 340, 2008 WL 131946; Rev. Rul. 2005-40, 2005-2 C.B. 4, 2005 WL 1415557; Rev. Rul. 2001-31, 2001-1 C.B. 1348, 2001 WL 606232; Rev. Rul. 78-338, 1978-2 C.B. 107, 1978 WL 41909; Rev. Rul. 77-316, 1977-2 C.B. 53, 1977 WL 43573, *obsoleted by* Rev. Rul. 2001-31, 2001-1 C.B. 1348, 2001 WL 606232; I.R.S. Priv. Ltr. Rul. 96-24-028 (June 14, 1996), 1996 WL 326435.

58. *See* Christopherson, *supra* note 15, at 127–28; Gantt, *supra* note 15, at 517–20; Gomez, *supra* note 15, at 624–28; Knight & Knight, *supra* note 15, at 409; Minto, *supra* note 15, at 841–47; Schaller & Harshman, *supra* note 12, at 254–55; Singer, *supra* note 15, at 123, 133–34; Taylor, *supra* note 15 *passim*; Winslow, *supra* note 15, at 94–99.

59. *See, e.g.*, Rev. Rul. 2008-08, 2008-1 C.B. 340, 2008 WL 131946 (emphasizing that risk-shifting and risk-distribution must be present in order for an arrangement to constitute insurance for federal income tax purposes).

tims of the two-part "insurance" test.[60] To some extent, their bad luck was fortuitous because it resulted in a degree of healthy doctrinal modification[61] and, at the same time, provided an accessible setting for thorough doctrinal assessment.

Part II of this Article provides an in-depth examination of the rigorous attempt to apply the risk shifting/distributing standard in the captive insurance company context, the industry's creative responses, the conceptual concessions made by the Service, and the structural limits imposed on taxpayers. Part II closes by presenting the doctrinal status quo and discussing two recent IRS pronouncements that clarify the Service's current litigating position and serve as analytical segues to the critique presented in Part III and proposals presented in Part IV.

II. DOCTRINAL EVOLUTION AND THE STATUS QUO: CAPTIVE INSURANCE COMPANIES

A. Captive Insurance Companies—General Description and the Lurking Problem

Although it is easy to think of captive insurance companies existing only as wholly-owned domestic subsidiaries providing coverage to their parent or brother-sister entities (a "pure" captive), in reality captive insurance companies take many forms.[62] In addition to those organized in the United States, pure captives are commonly organized offshore in foreign jurisdictions.[63] Often this is because foreign jurisdictions have less demanding standards with respect to insurance company formation

60. *See, e.g.*, Malone & Hyde, Inc. v. Comm'r, 62 F.3d 835, 840 (6th Cir. 1995) (holding that "insurance" cannot exist between a "sham" captive and insureds in its corporate family); Carnation Co. v. Comm'r, 640 F.2d 1010, 1012–13 (9th Cir. 1981) (holding that no "insurance" exists if a company's risks are purportedly insured with an unrelated entity which reinsures those risks with the insured's wholly-owned insurance subsidiary); Rev. Rul. 77-316, 1977-2 C.B. 53, 1977 WL 43573 (setting forth the "economic family" theory as justification for denying deductions with respect to amounts paid to captive insurance companies—i.e., those within the same corporate family—because such amounts did not constitute "insurance" premiums), *obsoleted by* Rev. Rul. 2001-31, 2001-1 C.B. 1348, 2001 WL 606232.

61. *See, e.g.*, Amerco, Inc. v. Comm'r, 979 F.2d 162, 166, 168 (9th Cir. 1992) (holding that a parent, and a brother–sister entity, can have a true insurance transaction with the parent's captive insurance company when the captive has substantial outside insurance business); Harper Group v. Comm'r, 979 F.2d 1341, 1342 (9th Cir. 1992) (holding that a captive insurance company with 29%–33% of its business coming from unrelated insureds could have true "insurance" relationships with members of its corporate family); Sears, Roebuck & Co. v. Comm'r, 972 F.2d 858, 860, 864 (7th Cir. 1992) (holding that Sears purchased "insurance" from its wholly-owned subsidiary, Allstate, and noting that 99.75% of Allstate's business came from other sources); Humana, Inc. v. Comm'r, 881 F.2d 247, 257 (6th Cir. 1989) (emphasizing the *Moline Properties* mandate and holding that there was risk-shifting and risk-distributing when a captive insured the risks of brother–sister entities); Rev. Rul. 2001-31, 2001-1 C.B. 1348, 2001 WL 606232 (noting that the Service "will no longer raise the 'economic family theory,'" set forth in Rev. Rul. 77-316, "in addressing whether captive insurance transactions constitute valid insurance" and indicating that "the Service will address such transactions on a case-by-case basis").

62. *See, e.g.*, Minto, *supra* note 15, at 841–42.

63. *See* Schaller & Harshman, *supra* note 12, at 252.

and lower levels of ongoing oversight.[64] Several businesses may come together to form a "group captive" (which insures the various liability risks of its owners/members),[65] and it is not uncommon for a given business (whose clients routinely need to secure some form of insurance) to form a "sponsored captive"[66] for the benefit of its clients. In that setting, the sponsored captive creates segregated financial cells to isolate and thereby protect individual client risks from the risks of others.[67]

Whatever the form, companies generally resort to a captive insurance arrangement to secure coverage that is either unavailable or difficult to obtain at a reasonable price (for example, insurance for medical malpractice,[68] latent construction defects, or potential earthquake damage[69]). Alternatively, companies may also resort to a captive for financial or strategic reasons,[70] such as to lower the cost of insurance or invest the insurance company's reserves.[71] These legitimate, rational business purposes have never managed, on their own, to establish the existence of an "insurance arrangement"[72] under the *Le Gierse* standards.

Indeed, captive insurance company arrangements remain consistently vulnerable to attack because a company purportedly paying a premium to the captive bears an uncomfortable resemblance to a company contributing to a contingency reserve, and the Service has long been hostile to the notion that a taxpayer deserves a deduction for *that*.[73] For example, in *Spring Canyon Coal Co. v. Commissioner*,[74] the taxpayer operated a coal mine and was obligated by state law to obtain workmen's compensation insurance either by securing it from a private carrier, participating in the state insurance fund, or by demonstrating an independent ability to pay claims.[75] Opting to take the latter route, the taxpayer estab-

---

64. *See id.* at 253–54; *see also* Christopherson, *supra* note 15, at 132–33 (pointing out that offshore formation of captives is common because offshore domiciles have lower capitalization requirements, more lax reporting requirements, and a less stringent regulatory environment).

65. *See* Minto, *supra* note 15, at 842.

66. *See id.*

67. *Id.*

68. *See* Gantt, *supra* note 15, at 498–500 (noting that medical malpractice insurance needs have prompted many responses, including the formation of captive insurance companies).

69. *See* Schaller & Harshman, *supra* note 12, at 252.

70. *See* Christopherson, *supra* note 15, at 122 (stating that, among other reasons, health care providers may form a captive to control premium and capital investment and obtain coverage broader than that which is commercially available). Christopherson also notes that by forming a captive, a company should be able to minimize administrative costs, eliminate third-party profit margins, reduce loss exposures, and improve claim management. *See id.* at 123.

71. *See* Schaller & Harshman, *supra* note 12, at 252.

72. *See* Beech Aircraft Corp. v. United States, 797 F.2d 920, 923 (10th Cir. 1986) (emphasizing that the legitimacy of the business purpose behind forming the captive does not establish risk-shifting and risk-distributing and does not foreclose a finding of the absence of either).

73. *See* Gulf Oil Corp. v. Comm'r, 89 T.C. 1010, 1024 (1987) ("Although technically transfer of risk may occur when a captive is involved that is a separate, viable entity, financially capable of meeting its obligations, we simply declined to recognize it as such when the arrangement was merely in substance the equivalent of a reserve for losses or self-insurance.").

74. 43 F.2d 78 (10th Cir. 1930).

75. *Id.* at 78.

lished a "Welfare or Compensation Insurance Fund"[76] and ultimately sought to deduct amounts contributed to the fund as ordinary and necessary business expenses.[77] The Service objected, claiming the amounts were not "expenses" but rather "reserves" set aside for contingent losses.[78] Siding with the Commissioner, the Court of Appeals for the Tenth Circuit held that:

> The whole object of self-insurance is to avoid the expense of insurance premiums. If the petitioner had elected to insure its risks in the state fund or a private company, it would have expended the premium and shifted the risk; instead, it retained the risk and kept the premium. Having elected not to expend the premium, it cannot charge a corresponding sum as an "expense." . . . [T]he petitioner is not entitled to deduct as an expense a sum of money which it might have expended for insurance premiums, but did not.[79]

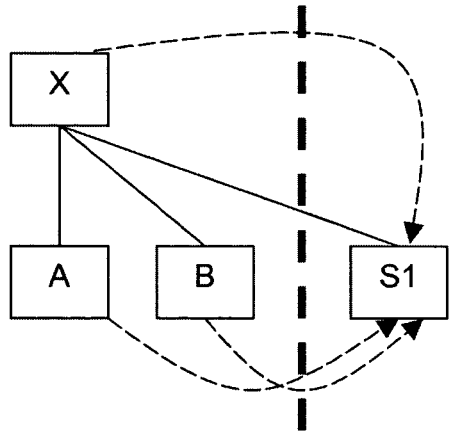
The dilemma faced by taxpayers in situations like those in *Spring Canyon Coal Company* did not resolve itself over time. Aggressively seeking (or legally obligated) to acquire some form of designated insurance, and determined to do so in a manner they deemed financially feasible, companies were rather creative in their efforts to get both the coverage and the benefit of something akin to a premium payment deduction.[80] The Service did not sit idly by. To the extent that companies steered clear of naked contingency fund reserving and moved towards the use of separately-incorporated but wholly-owned "insurance" entities, the Service countered by creating or appealing to broad-sweeping doctrines to disallow deductions for what it considered contingency reserve contributions.[81] For decades, the "economic family" theory was the Service's weapon of choice.[82]

B. The Rise and Fall of the "Economic Family" Argument Empire

The Service formally set forth its "economic family" theory in Revenue Ruling 77-316[83] by presenting and assessing three hypothetical situations. In each situation, a new, wholly-owned foreign insurance subsidiary provides fire and other casualty insurance to members of its corporate family.

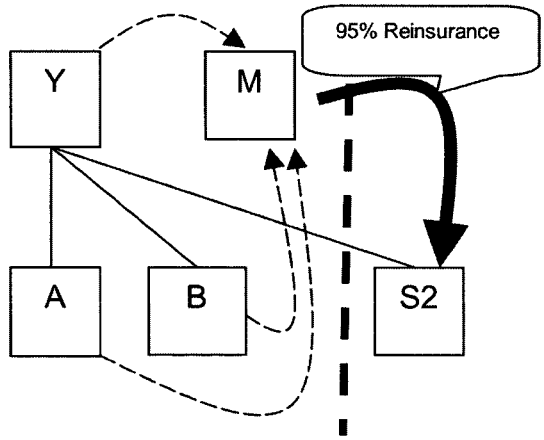
---

76. *Id.* at 79.
77. *Id.*
78. *Id.*
79. *Id.* at 80.
80. *See, e.g.*, Carnation Co. v. Comm'r, 640 F.2d 1010, 1013 (9th Cir. 1981) (using a fronting company obligated to reinsure with a wholly-owned captive insurance subsidiary and using a contingent capitalization agreement to persuade the fronting company).
81. *See* Rev. Rul. 77-316, 1977-2 C.B. 53, 1977 WL 43573 (announcing the "economic family" theory), *obsoleted by* Rev. Rul. 2001-31, 2001-1 C.B. 1348, 2001 WL 606232.
82. *See, e.g., id.*; Sedore, *supra* note 15, at 1105. Commentators note that the Service has challenged the current deductibility of premiums paid to captive insurance companies by appealing to various theories. *See, e.g.*, Sedore, *supra* note 15, at 1105–06.
83. *See* Rev. Rul. 77-316.



REV. RUL. 77-316, SITUATION 1

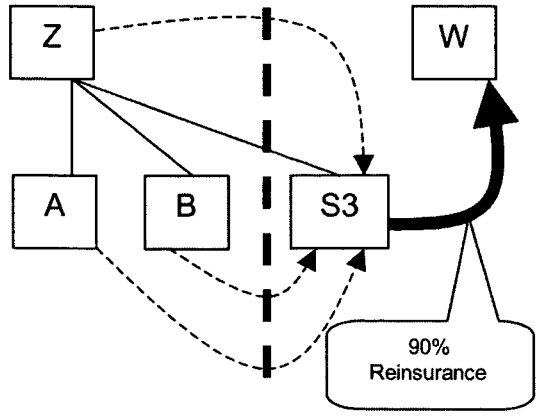
In *Situation 1*, X and its wholly-owned domestic subsidiaries, A and B, paid amounts to S1, a wholly-owned foreign "insurance" subsidiary of X, for various forms of property and casualty insurance. The rates paid by X, A, and B were commercially appropriate, and S1 provided insurance only to X, A, and B.



REV. RUL. 77-316, SITUATION 2

In *Situation 2*, the scenario is similar, however Y and its wholly-owned subsidiaries, A and B, paid amounts to M, an unrelated insurance company (a "fronting company"[84]) which, by agreement, reinsured 95% of this risk with S2, Y's wholly-owned foreign insurance subsidiary.

84. *See* Minto, *supra* note 15, at 847.



REV. RUL. 77-316, SITUATION 3

In *Situation 3*, the facts are similar to *Situation 1*, but Z and its wholly-owned, domestic subsidiaries, A and B, paid amounts to S3, and by prior agreement, S3 reinsured 90% of this risk with W, an unrelated insurance company.

In assessing the deductibility of amounts paid from the parent company and subsidiaries to the relevant entity, Revenue Ruling 77-316 starts by emphasizing both the risk shifting/distributing standards set forth in *Le Gierse* and the historical judicial hostility directed towards taxpayer attempts to take deductions for contributions to self-insurance reserves.[85] The Service concluded that the amounts paid over to S1, S2, and S3 do not involve risk shifting or risk distributing because the parent company, insured subsidiaries, and captive all "represent[ed] one economic family."[86] Accordingly, the Service reasoned that those suffering the loss ultimately bear the financial burden of the loss.[87] The Service did clarify that an arrangement constituted insurance to the extent risks were retained by third-party insurers or passed on to reinsurers (for example, the 5% retained by the fronting company in *Situation 2* and the 90% passed on in *Situation 3*).[88] Further, amounts paid by the purported insured subsidiaries (as premiums) and amounts paid by the captive (as loss coverage) to those subsidiaries were generally recast as dividends from one entity to the common parent followed by a contribution of a given amount to the intended recipient as a capital contribution; "premium" payments from the parent to the captive were recast as capital contributions to the captive, and loss coverage payments from the captive to the parent were deemed to be dividends to the parent.[89] Given that its conclusions were facially at odds with *Moline Properties, Inc. v. Com-*

85. Rev. Rul. 77-316.
86. *Id.*
87. *Id.*
88. *See id.*
89. *See id.*

missioner,[90] which dictates that an entity organized for legitimate business activity is treated as a separate taxable entity, the Service attempted to reconcile the differences. The Service emphasized that it respected the separate corporate existence of the various entities (giving due regard to their business activity), but opted to "examine[] the economic reality" presented by each hypothetical situation.[91]

For several years, the Service's economic family argument managed to secure a degree of traction, especially in the U.S. Court of Appeals for the Ninth Circuit. In *Carnation Co. v. Commissioner*, the taxpayer created a wholly-owned Bermuda subsidiary to insure its risks, as well as the risks of several of its subsidiaries.[92] By prior arrangement, Carnation paid premiums directly to American Home Assurance Company, which proceeded to reinsure 90% of Carnation's business with Carnation's captive insurance subsidiary. It was clear, however, that American Home would not have entered into an agreement with the captive unless Carnation agreed to further capitalize the captive if needed.[93] Denying the deductibility of 90% of the premium paid to American Home, the court refused to countenance the fronting company strategy adopted by Carnation.[94] It also affirmatively rejected any *Moline Properties* arguments,[95] and thereby substantially aligned itself with the Service's position in *Situation 2* in Rev. Rul. 77-316.

A few years later, the Ninth Circuit drove the point home in *Cloughtery Packing Co. v. Commissioner*.[96] *Cloughtery Packing* held that, even in the absence of a contingent capitalization agreement, a parent corporation's payment of premiums to an unrelated insurer who, by pre-arrangement, reinsures that risk with a captive insurance company of the parent is not payment of an "insurance" premium because the parent "retains an economic stake in whether a covered loss occurs."[97] All along, it should have been known that the outright defiance of the *Moline Properties* mandate via the economic family theory could not last.

---

90. 319 U.S. 436, 438–39 (1943). ("The doctrine of corporate entity fills a useful purpose in business life. Whether the purpose be to gain an advantage under the law of the state of incorporation or to avoid or comply with the demands of creditors or to serve the creator's personal or undisclosed convenience, so long as that purpose is the equivalent of business activity or is followed by the carrying on of business by the corporation, the corporation remains a separate taxable entity.").

91. Rev. Rul. 77-316.

92. Carnation Co. v. Comm'r, 640 F.2d 1010, 1012 (9th Cir. 1981).

93. *See id.*

94. *See id.* at 1013.

95. *Id.* ("We reject Carnation's contention that [language quoted from Revenue Ruling 77-316] conflicts with the recognition of the separate status of corporations.").

96. 811 F.2d 1297 (9th Cir. 1987).

97. *Id.* at 1299, 1307 (indicating that there was no indemnification, capitalization, or other guarantee in place with respect to the parent company and the wholly-owned captive insurance company).

Although the captives in *Carnation* and *Clougherty Packing* insured no one outside their affiliated group,[98] a number of wholly-owned or controlled insurance companies (which are technically captive insurance companies relative to their parent entities) have a substantial amount of outside business.[99] To conclude that such "captives" can have valid insurance relationships with non-affiliated entities, while being unable to have such relationships with affiliated entities, would defy logic and effectively gut *Moline Properties*. The Tax Court embraced this reality in *Gulf Oil Corp. v. Commissioner*.[100] Even though the court concluded that the percentage of insurance business from unrelated parties was insufficient,[101] it noted that the existence of sufficient outside business would clear the self-insurance taint of an arrangement between various affiliated entities and their captive insurance company.[102]

Several years later, the U.S. Court of Appeals for the Seventh Circuit adopted this view. In *Sears, Roebuck & Co. v. Commissioner*, the court held that Sears was purchasing true "insurance" from its wholly-owned subsidiary, Allstate,[103] because 99.75% of Allstate's business came from other sources.[104] Other Courts of Appeals, as well as the Service, have since clarified that lower percentages of outside business will suffice,[105] but the more significant development lies in the fact that the decisions collectively establish the relative irrelevance of risk shifting in the face of substantial risk distribution. Even in the absence of a risk-shifting imperative, however, the Service could still require the existence of substantial outside business before acquiescing to the existence of a valid insurance relationship[106]—or so it thought.

Even before *Sears, Roebuck and Co.* and its progeny challenged the core legitimacy of a risk shifting inquiry, at least one court had begun to

---

98. *Clougherty Packing Co.*, 811 F.2d at 1299; *Carnation Co.*, 640 F.2d at 1012.

99. *See, e.g.*, Sears, Roebuck & Co. v. Comm'r, 972 F.2d 858, 860 (7th Cir. 1992) (noting that Allstate, a subsidiary of Sears, had substantial outside business).

100. 89 T.C. 1010, 1027–28 (1987) (concluding that the fact that the reinsuring captive insurance company had 2% net premium income from unrelated parties did not result in the characterization of "premiums" paid indirectly by the parent corporation and its affiliates as "insurance" premiums).

101. *Id.*

102. *Id.* at 1026–27.

103. 972 F.2d 858, 864 (7th Cir. 1992).

104. *Id.* at 860.

105. *See, e.g.*, Ocean Drilling & Exploration Co. v. United States, 988 F.2d 1135, 1152 (Fed. Cir. 1993) (holding that true "insurance" existed between a captive insurance company and the insureds in its corporate family when the captive had 44%–66% of its business coming from unrelated entities); Amerco, Inc. v. Comm'r, 979 F.2d 162, 168 (9th Cir. 1992) (finding sufficient unrelated business at 52%–74% of total business); Harper Group v. Comm'r, 979 F.2d 1341, 1342 (9th Cir. 1992) (citing unrelated business was 29%–33% percent of total business).

106. *See e.g.*, Gulf Oil Corp. v. Comm'r, 89 T.C. 1010, 1027–28 (1987) (concluding that 2% of net premiums from outside sources was sufficient to create an insurance relationship between a captive and its affiliated); *see also* Christopherson, *supra* note 15, at 128 (pointing out that courts will look to the level of a subsidiary's outside business in determining whether an insurance relationship exists between the parent and the subsidiary).

whittle away at the outside business requirement. In *Humana, Inc. v. Commissioner*, the U.S. Court of Appeals for the Sixth Circuit held, over the contrary opinion of the Tax Court, that risk shifting and risk distributing were present when a captive insurance company insured the risks of its brother-sister entities.[107] However, it did not extend to the parent-captive relationship.[108] The Tax Court, in considering the parent-captive relationship, concluded that the payment of a covered loss by the captive to the parent would ultimately show up on the parent's balance sheet as a reduction in the value of the captive's stock,[109] a result it found offensive in light of its belief that true risk shifting must occur and that "[t]rue insurance relieves the firm's balance sheet of any potential impact of the financial consequences of the insured peril."[110]

Although the Tax Court's logic could not apply with respect to payments for the losses of brother-sister entities of the captive, the court chose not to invite or sanction the favorable manipulation of corporate structure.[111] Emphasizing the *Moline Properties* mandate[112] and firmly rejecting the Service's economic family argument,[113] the Sixth Circuit opted to go a different route, respecting the balance sheet analysis with regard to the parent, but finding a valid insurance relationship between the brother-sister entities and the captive. Other courts soon followed suit.[114]

In addition to weathering judicial assault, the economic family theory found itself the target of scholarly criticism. One commentator charged the Service with abusing its discretion by inventing and advocating an "eccentric theory of law" to "provide a quick fix for a perceived abuse,"[115] rather than doing the legwork to mount an effective attack on more established grounds.[116] Although the commentator acknowledged

---

107. 881 F.2d 247, 257 (6th Cir. 1989) (emphasizing the *Moline Properties* mandate and holding that there was risk-shifting and risk-distributing when a captive insured the risks of brother–sister entities).

108. *Id.*

109. Humana, Inc. v. Comm'r, 88 T.C. 197, 211–12 & n.13 (1987).

110. *Id.* at 211.

111. *Id.* at 213–14.

112. *Humana, Inc.*, 881 F.2d at 252.

113. *Id.* at 257.

114. *See, e.g.*, Hosp. Corp. of Am. & Subsidiaries v. Comm'r, 74 T.C.M (CCH) 1020, 1997 WL 663283, at *18, *26–27 (1997) (concluding that, between a captive insurance subsidiary and its brother–sister entities, there was bona fide "insurance"); Kidde Indus., Inc. v. United States, 40 Fed. Cl. 42, 67 (1997) (holding that a parent company could deduct amounts paid—to cover its separately-incorporated subsidiaries—to unrelated insurer ceding amounts to parent's captive insurance subsidiary but denying such treatment with respect to payments made on behalf of the parent's divisions), *appeal dismissed per stipulation*, 194 F.3d 1330 (Fed. Cir. 1999) (unpublished table decision).

115. Singer, *supra* note 15, at 115.

116. *See id.* at 127 (arguing that the Service might have been able to attack the transactions by using § 482). Singer also notes that the Service may have had little confidence in existing precedent to support a transaction recharacterization approach. *See id.* at 123.

that taxpayers might form captives in attempts to abuse the tax system,[117] he went on to endorse an alternative that would have the Service focus on the captive (i.e., the income side) and require that it satisfy certain requirements in order to qualify for special treatment.[118]

Other commentators, like some courts, highlighted the simple inability to reconcile the economic family theory with the mandate of *Moline Properties*.[119] Even while noting that companies forming captive insurance companies may be motivated, to some extent, by tax considerations, they reasoned that such motivations have never sufficed to disallow premium deductions at the parent level, especially in light of the legitimate business reasons for forming captives[120] and the doctrine that a subsidiary be treated as a separate and distinct entity when the parent effectively establishes the subsidiary's business purpose.[121]

In response to the occasional irrelevance of a risk-shifting inquiry,[122] ready judicial acceptance of brother-sister-captive insurance relationships,[123] scholarly criticism,[124] and overt judicial rejection of the economic family argument,[125] the Service issued Revenue Ruling 2001-31.[126] In doing so, the Service promised that it would "no longer raise the 'economic family theory' set forth in Revenue Ruling 77-316 . . . in addressing whether captive insurance transactions constitute valid insurance."[127] The Service did emphasize that, in general, it would continue to analyze individual arrangements on a case-by-case basis and specifically challenge some arrangements, such as alleged sham transactions[128] and

117. *Id.* at 117.

118. *Id.* at 162–64. Singer would generally require proper record keeping and management, adequate and properly invested reserves, proper risk pooling, arms-length pricing, and maintenance of a sufficient level of unrelated policyholder business. *Id.* at 163. Singer also set forth simple disallowance of a deduction for premiums paid to a related party, but he questioned the wisdom of that approach, given the ordinariness of transferring risk to a legitimate insurance company. *Id.* at 162.

119. Knight & Knight, *supra* note 15, at 417.

120. *Id.* at 404.

121. *Id.*

122. *See, e.g.*, Sears, Roebuck & Co. v. Comm'r, 972 F.2d 858, 861 (7th Cir. 1992).

123. *See, e.g.*, Humana, Inc. v. Comm'r, 881 F.2d 247, 255 (6th Cir. 1989).

124. *See, e.g.*, Christopherson, *supra* note 15, at 128 (noting that the "economic family approach has been widely rejected"); Gomez, *supra* note 15, at 627–28 (arguing for the eradication of the economic family theory and the institution of a two-part inquiry into the legitimacy of the insurance company and the particulars of given transactions).

125. *See Humana, Inc.*, 881 F.2d at 257.

126. Rev. Rul. 2001-31, 2001-1 C.B. 1348, 2001 WL 606232.

127. *Id.* (internal citation omitted).

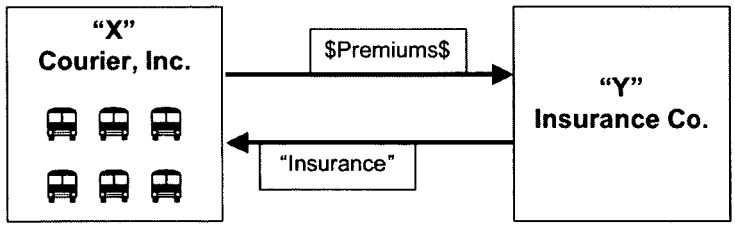
128. The Service specifically referenced *Malone & Hyde, Inc. v. Commissioner*, which held that "insurance" cannot exist between a "sham" captive and insureds in its corporate family. *Id.* (citing Malone & Hyde, Inc. v. Comm'r, 62 F.3d 835 (6th Cir. 1995)). According to the court, a captive will be a "sham" captive if it is either (1) undercapitalized, (2) a party to a reinsurance arrangement in which the captive's parent indemnifies the primary insurer for amounts owed by the captive to the primary insurer, or (3) both. *Malone & Hyde, Inc.*, 62 F.3d at 842–43.

purported insurance transactions between parent entities and their captives.[129]

C. The Status Quo

Even though the Service officially abandoned its economic family argument in Revenue Ruling 2001-31, it tenaciously adheres to and applies the risk shifting/distributing standard in assessing purported insurance arrangements.[130] This may well be, in some contexts, a means of appealing to the economic family theory under a safer rubric. Indeed, recent IRS pronouncements demonstrate that the Service has already taken the entire game to the next level. Rather than merely assessing the existence of an insurance relationship when the transaction involves affiliated entities, the Service is now willing to scrutinize transactions between purported insureds and entities *wholly unaffiliated* with the insured or its group (even when there is no reinsurance with a controlled subsidiary of the insureds). Arguably, this move represents the opening salvo of a new wave of intent: the elevation of form over substance in the captive arena.

In Revenue Ruling 2005-40,[131] the Service presented and assessed four situations to clarify its current litigating position. Each situation is presented schematically with the Service's assessment appearing immediately thereafter. All situations involve a domestic entity, X, which operates a courier transport business, owns a substantial fleet of automotive vehicles, needs to insure those vehicles against various operational risks, and pays an arm's length premium to Y for the coverage. Note that in each instance, X owns vehicles "representing a significant volume of independent, homogenous risks"[132] (that is, enough to distribute risk independently), and that Y is adequately capitalized and operates in accord with state law.[133]



REV. RUL. 2005-40, SITUATION 1

129. *See, e.g.*, Clougherty Packing Co. v. Comm'r, 811 F.2d 1297, 1307 (9th Cir. 1987) (holding that a parent–subsidiary transaction was not insurance).

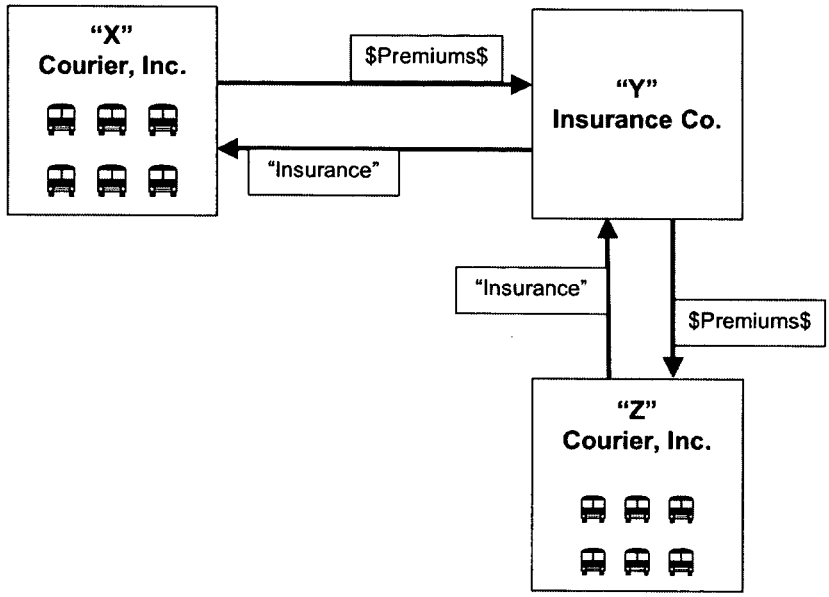
130. Rev. Rul. 2008-08, 2008-1 C.B. 340, 2008 WL 131946; Rev. Rul. 2005-40, 2005-2 C.B. 4, 2005 WL 1415557.

131. Rev. Rul. 2005-40.

132. *Id.*

133. *Id.*

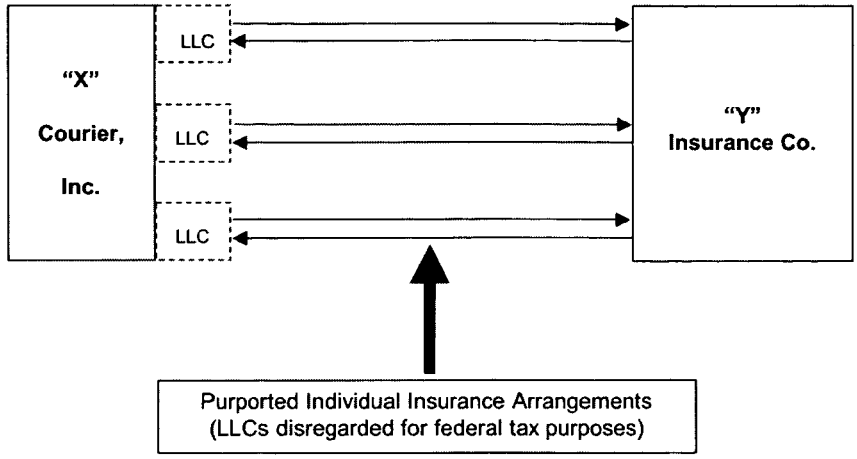
In *Situation 1*, the X–Y contract constitutes 100% of Y's business. The Service acknowledges that X shifted risk from itself to Y, but, after noting that Y does not insure any non-X risks, concludes that Y has not distributed X risks amongst other policyholders. Accordingly, the Service concludes that the relationship between X and Y is not an insurance relationship.[134]



REV. RUL. 2005-40, SITUATION 2

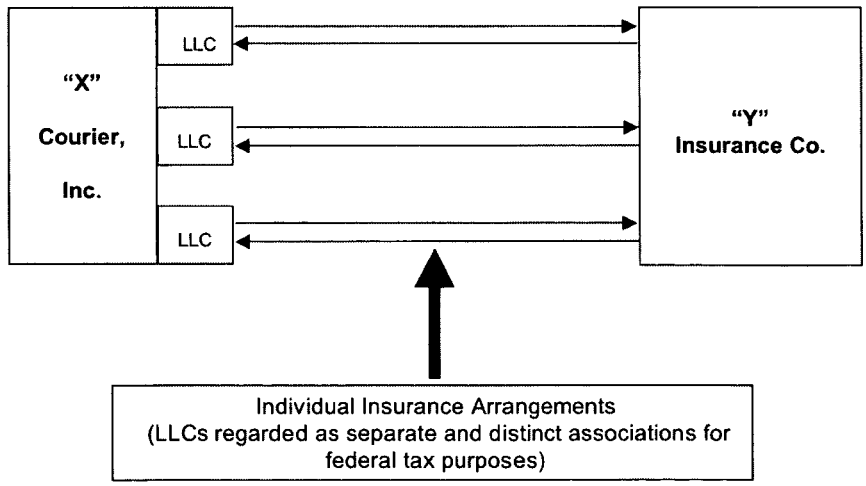
The facts in *Situation 2* indicate that the X–Y contract constitutes 90% of Y's business and that the remaining 10% comes from Y's contract with Z. Here too, the Service concludes that the relationship between X and Y is not an insurance relationship. Although Y's relationship with Z adds some degree of risk distribution, the Service concludes that the 10% of risk from Z is insufficient to alter the character of the X–Y relationship.[135] Therefore, the relationship does not constitute "insurance."

134. *Id.*
135. *Id.*



REV. RUL. 2005- 40, SITUATION 3

In *Situation 3*, Y does not contract with X. Instead, Y contracts with twelve LLCs, each of which has X as its single member. X conducts its business through the LLCs, and each LLC has elected to be disregarded for federal income tax purposes (that is, each LLC has elected not to be treated as separate and distinct from X). The facts also indicate that each contract between Y and a disregarded LLC constitutes 5%–15% of Y's total business and that Y insures no other entities. Because Y is treated as contracting only with X, the Service concludes that the individual arrangements do not constitute insurance for the same reasons set forth in *Situation 1*.[136]



REV. RUL. 2005 - 40, SITUATION 4

Situation 4 is the same as *Situation 3*, except that the LLCs elect to be treated as entities separate and distinct from X. Under these facts, the

136. *Id.*

Service concludes that the individual LLCs shift risk to Y and that the risk is distributed amongst the twelve LLCs. Accordingly, the relationships are deemed to constitute "insurance" relationships.[137]

The Service's most recent pronouncement in the captive insurance arena is Revenue Ruling 2008-08,[138] which involves a sponsored captive in the form of a "protected cell company."[139] A protected cell company is similar, in many ways, to a group captive, an arrangement in which various companies come together and form a captive for the purpose of insuring their own risks.[140] The Service has ruled that premium payments to a group captive owned by thirty-one companies were "insurance" premiums,[141] noting that "because the taxpayer and the other insureds-shareholders are not economically related, the economic risk of loss can be shifted and distributed among the shareholders who comprise the insured group."[142] A protected cell company differs from a group captive in that separate cells/accounts are created for each equity holder making a capital contribution, and each cell receives premiums and insures the risks of the designated equity holder.[143] Revenue Ruling 2008-08 clarifies that insurance may or may not exist with respect to a given protected cell company relationship.

137. *Id.*

138. Rev. Rul. 2008-08, 2008-1 C.B. 340, 2008 WL 131946.

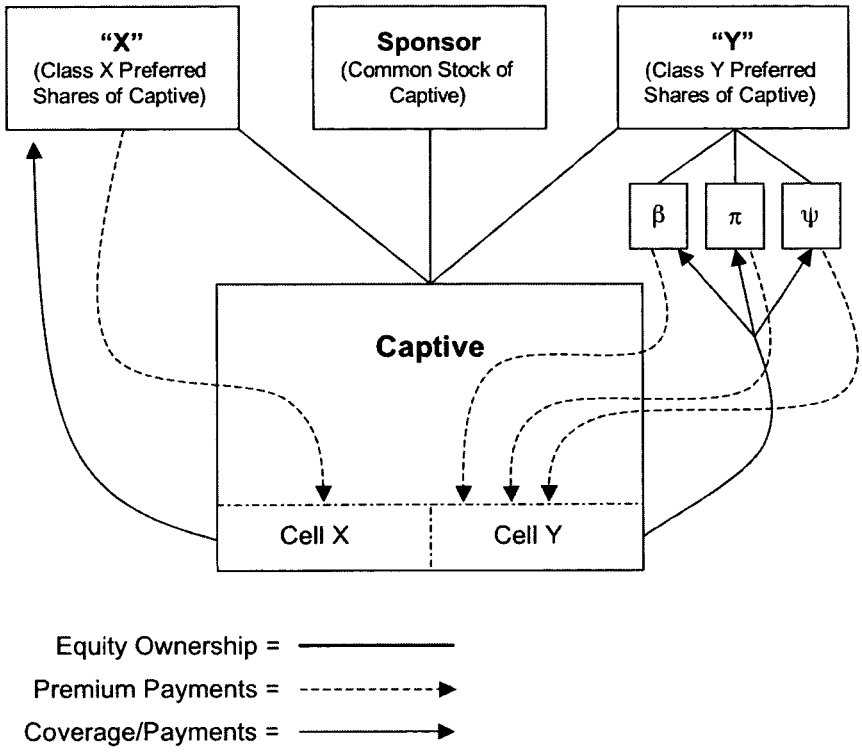
139. *Id.*

140. *See* Minto, *supra* note 15, at 842.

141. Rev. Rul. 78-338, 1978-2 C.B. 107, 1978 WL 41909 (concluding that premium payments made to a group captive owned by 31 companies were ordinary and necessary business expenses—i.e., "insurance" premiums).

142. *Id.*

143. Gantt, *supra* note 15, at 520 ("With the protected cell, investor funds are not subject to losses affecting other parts of an insurance company. Monies from investors would be deposited into a protected cell and used for a specific delineated purpose, such as catastrophic loss.").



REV. RUL. 2008-08

In this protected cell company arrangement, X only pays premiums to Cell X, and Cell X only insures the risks (and pays the losses) of X.[144] Similarly, Cell Y only receives payments from β, π, and ψ and only insures the risks (and pays the losses) of those entities.[145] The Service concluded that the arrangement between Cell X and X does not constitute an insurance arrangement because the relationship is akin to one in which a subsidiary insures only its parent.[146] Accordingly, risk shifting and risk distributing are not present; thus, no insurance relationship exists.[147] Noting that the relationship between Cell Y and the wholly-owned subsidiaries of Y were akin to a brother-sister insurance relationship, the Service concluded that the arrangements constituted insurance given the presence of adequate risk shifting and risk distributing.[148]

III. CRITIQUE

Although the Service officially asserted that it would no longer employ the "economic family" argument in assessing captive insurance

144. Rev. Rul. 2008-08, 2008-1 C.B. 340, 2008 WL 131946.
145. *Id.*
146. *Id.*
147. *Id.*
148. *Id.*

company arrangements,[149] recent IRS pronouncements demonstrate it did not tell the whole truth. The prevailing common law environment will not allow an economic family argument to succeed in a captive-brother-sister setting.[150] But, depending on the level of non-parent business at the subsidiary level, the Service can still get away with it in some parent-captive contexts (or those deemed akin to that context) so long as the position is cloaked in the garb of risk shifting and risk distributing.[151] Assuming *arguendo* that the risk shifting and risk distributing standards were ever truly useful, they now arguably fall short as effective determinants, especially in today's complex business environment.

A. The Shell of Risk Shifting

Sears, Roebuck & Co. and its progeny have firmly established that an insurance relationship exists between a parent company and its wholly-owned subsidiary when the subsidiary has a sufficient level of outside business.[152] Thus, although the Service routinely cites *Le Gierse* on the way to its risk shifting and risk distributing analyses,[153] various courts (whether explicitly or implicitly) fail to regard risk shifting, in the traditionally-understood sense, as a *sine qua non* of insurance.[154] To the extent a sufficient level of risk distribution is the key to eliminating the risk shifting requirement, established companies with substantial outside business enjoy an undue benefit relative to new and growing companies.

Although a given company may, after further development, reap the same benefits, there remains the troubling problem of uni-directionality. Under *Sears, Roebuck & Co.*, an insurance relationship is deemed to exist between a parent company and its wholly-owned captive insurance subsidiary because the captive has a sufficient level of outside business to distribute its risks adequately (in other words, because the *captive* has enough homogenous, independent risks to satisfy the law of large numbers).[155] This reality does not change when the parent company adds its risks to the mix, but, per the confirmation of Revenue Ruling 2008-08, the street only runs one way.[156] If it is the *parent* (or a parent-like entity) with enough homogenous, independent risks to satisfy the law of large numbers, the Service will not respect the parent-captive arrangement as

149. Rev. Rul. 2001-31, 2001-1 C.B. 1348, 2001 WL 606232.

150. *See, e.g.*, Humana, Inc. v. Comm'r, 881 F.2d 247, 248–52 (6th Cir. 1989).

151. *See, e.g.*, Hosp. Corp. of Am. & Subsidiaries v. Comm'r, 74 T.C.M. (CCH) 1020, 1997 WL 663283, at *24–*26 (1997); Kidde Indus., Inc. v. United States, 40 Fed. Cl. 42, 51–57 (1997), *appeal dismissed per stipulation*, 194 F.3d 1330 (Fed. Cir. 1999) (unpublished table decision).

152. *See, e.g.*, cases cited *supra* note 61.

153. *See, e.g.*, Rev. Rul. 2008-08, 2008-1 C.B. 340, 2008 WL 131946.

154. *See, e.g.*, *Humana, Inc.*, 881 F.2d at 252–56.

155. Sears, Roebuck & Co. v. Comm'r, 972 F.2d 858, 860 (7th Cir. 1992) (noting that Allstate collects more than $5 billion in premiums annually and that "[s]ome 99.75% of Allstate's premiums [come] from customers other than Sears").

156. Rev. Rul. 2008-08 (citing Rev. Rul. 2005-40, 2005-2 C.B. 4, 2005 WL 1415557; Rev. Rul. 2002-89, 2002-2 C.B. 984, 2002 WL 31749579) (indicating that the relationship between X and Cell X would not qualify as insurance).

"insurance" in the absence of outside business at the captive level, because risk shifting and risk distribution are deemed absent.[157] That is, the Service would refuse to find "insurance" despite the presence of effective risk distribution and arguable irrelevance of the risk shifting standard. A far more egregious problem arises when the Service elects to ignore risk-distributing realities in a similar context: when a *non-affiliated* insurer is involved and risk shifting cannot be denied.

B. The Husk of Risk Distribution: Doctrinal Issues

Under Revenue Ruling 2005-40, *Situation 1*, the Service will not respect an arrangement as insurance where a single entity with risk distributing capacity happens to be the only insured with respect to an *unrelated* entity.[158] Arguably, this approach represents an aggressive elevation of form over substance. X's homogenous, independent risk units are lumped together under a single corporate roof[159] such that, for risk distribution purposes, X is deemed to be nothing more than a single insured corporate entity.[160] And while the Service appears to acknowledge the existence of risk shifting,[161] its final conclusion on *Situation 1* reveals something altogether suspect. Rather than recognizing the business reality (that Y is insuring a large volume of homogenous, independent risks such as individual vehicles), the Service deems Y's business activity to be the attempt to insure one entity, X, even though Y, by contract, would likely have been on the hook with respect to an individual vehicle loss, and not solely on the loss of the entire fleet. The Service's posture in this regard is perplexing in light of the Tax Court's conclusion in *Gulf Oil Corp. v. Commissioner* that "a single insured can have sufficient unrelated risks to achieve adequate risk distribution,"[162] and the IRS's pronouncement indicating that "a single taxpayer may transfer an amount of homogenous and statistically independent risks which would be sufficient to satisfy the risk distribution requirement."[163]

The question then becomes whether it is appropriate to pierce the veil of the nominal insured for risk distribution assessment purposes.[164] Surprisingly, as discussed below, the Service willingly confirmed in at least one instance that individual independent risk exposures are meas-

157. *Compare* Rev. Rul. 2008-08 (finding that no insurance relationship existed in a given transaction because it was akin to a parent-subsidiary relationship), *with Sears, Roebuck & Co.*, 972 F.2d at 860 (7th Cir. 1992) (finding that insurance existed when a subsidiary with substantial outside business insured its parent).

158. Rev. Rul. 2005-40, 2005-2 C.B. 4, 2005 WL 1415557.

159. *Id.*

160. *Id.*

161. *See id.*

162. Gulf Oil Corp. v. Comm'r, 89 T.C. (CCH) 1010, 1026 (1997).

163. I.R.S. Field Serv. Adv. (Oct. 19, 1998), 1998 WL 34066011.

164. Logic and experience dictate that when an insurance company issues a group life insurance policy, there is no confusing insured individuals with the common corporate employer.

ured by looking *within* the corporate entity,[165] not by viewing the corporate entity as a single insured unit.

In Private Letter Ruling 9624028, the Service addressed the deductibility of amounts paid by thirty-six Funds to an assessable mutual insurance company owned solely by the thirty-six Funds and formed to insure them against various loss events with respect to the investments made by each Fund.[166] After highlighting the fact that none of the Funds owned a controlling interest in the mutual,[167] the Service noted the following:

> [I]n terms of risk distribution, we note that Mutual has accepted a large number of independent risks and is taking advantage of the "law of large numbers." (The (independent) risk exposures here are possible defaults of *any one* of the 28x issues of "securities" of Fund 1 (or the possible default of any of the issues of any other of the 35 funds).)[168]

Thus, the Service privately says that it will look beyond the corporate shell at individual risk units to ascertain whether risk has been adequately distributed, but pronounces something altogether different in *Situation 1* of Revenue Ruling 2005-40. The analytical consistency does not improve with *Situations 2–4* of that pronouncement.

In *Situation 2*, the Service concludes that the presence of an additional business source does not automatically alter the character of the X-Y relationship.[169] However, to the extent that this ruling and others designate or look to a specific percentage of outside business as key to the "insurance" analysis, they reveal nothing more than arbitrary line-drawing. The Service's conclusions regarding *Situations 3* and *4*, viewed together, would appear to allow, rather inexplicably,[170] a taxpayer to create or destroy "insurance" merely by making an election.[171] Yet, random elevations of form over substance have been prohibited in this arena since Edyth Le Gierse struck her deal with Connecticut General. Or at least that's been the long-running story.

C. The Husk of Risk Distribution: The Modern "Insurance" Reality

Wholly aside from doctrinal problems, practical realities do not jibe well with the Service's prevailing risk distribution standard. The canvas of background assumptions is too pristine. Although it may be comfort-

165. I.R.S. Priv. Ltr. Rul. 96-24-028 (June 14, 1996), 1996 WL 326435.
166. *Id.*
167. *Id.*
168. *Id.* (emphasis added).
169. Rev. Rul. 2005-40, 2005-2 C.B. 4, 2005 WL 1415557.
170. *See* Sedore, *supra* note 15, at 1104 (describing the Service's rationale in assessing *Situations 3* and *4* as "somewhat inexplicable" because in some instances, the Service respects the separate existence of disregarded LLCs).
171. *See id.*

ing to assume that insurance companies willingly offer life, health, and other forms of insurance to broad segments of the population and quietly allow the law of large numbers to operate, the truth is far nastier.

Modern insurance is a business more so than ever, and rather than suffer the vagaries of the law of large numbers, underwriters do their best to outsmart or outmaneuver it. Companies impose tough underwriting standards and thereby manage to issue unwarranted denials of coverage. And even when coverage is offered and premiums are paid over time in good faith, companies habitually attempt to deny coverage by appealing to, among other things, contractual exclusions for "pre-existing conditions" and "experimental" treatments.[172]

The case of Patrick Tumulty is typical. After paying health insurance premiums to Assurant Health for six years (under a series of six-month policies),[173] the company ultimately refused coverage for Tumulty's kidney disease.[174] Because the company treated him as a new insured under each contract,[175] his ailment was deemed a pre-existing condition when it was finally diagnosed[176]—prior medical tests under "prior" policies proved it.[177]

If certain commentators are to be believed, the problem of unwarranted coverage denial is rampant. In his searing documentary on the American health care system,[178] Michael Moore captured the prevailing sentiment. One claims adjuster noted, "'[People aren't] slipping through the cracks,' . . . '[The insurance companies] made the crack and are sweeping you toward it.'"[179] Thus, at least with respect to some forms of insurance, the theoretical operation of risk distribution stands in stark contrast to practical realities.

Further, at the same time that companies are attempting to minimize or eliminate the negative impact of the law of large numbers in one sphere, they are also flouting, misinterpreting, or simply ignoring the fundamental rules of traditional risk distribution elsewhere. As the following discussion demonstrates, the AIG saga adds an exclamation point to that assertion and helps reveal the truly limited utility of the risk distribution standard, as currently applied as an "insurance" determinant.

---

172. *See* Karen Tumulty, *The Health Care Crisis Hits Home*, TIME, Mar. 16, 2009, at 26, *available at* 2009 WLNR 4231105 (discussing the misfortune of Patrick Tumulty whose kidney disease was deemed to be a non-covered, pre-existing condition by Assurant Health, which issued him a series of six-month policies and treated him as a new insured under each contract).

173. *Id.*

174. *Id.*

175. *Id.*

176. *Id.*

177. *Id.*

178. SICKO (Weinstein Company 2007).

179. Richard Corliss, Sicko *Is Socko*, TIME.COM, May 19, 2007, http://www.time.com/time/arts/article/0,8599,1623337,00.html.

As noted earlier, risk distribution relies on the operation of the law of large numbers. With a sufficiently large pool of individuals, for example, actual mortality and predicted mortality will roughly correspond due to the accuracy of historical mortality and morbidity data. Thus, the law of large numbers operates well in some insurance environments (such as life insurance). If, however, a company seeks to insure against the occurrence of an event when there is no reliable data with respect to prior events, accurate probabilities cannot be ascertained, and the law of large numbers (no matter how large the pool of insureds) cannot be relied on to match predicted and actual experience going forward.[180] As explained below, AIG's treacherous dalliance with credit default swaps and special purpose vehicles bears this out.

D. Special Purpose Vehicles, Credit Default Swaps, and Sub-Prime Mortgages

Several years ago, there was a sharp and noticeable uptick in the making of subprime loans.[181] Certain existing and aspiring homeowners were easy targets.[182] Once originated, the individual mortgage obligations were transferred through the hands of various industry participants, and ultimately pooled together in a trust or other entity—such as a special purpose vehicle or "SPV"[183]—such that interested parties could invest in the pool by purchasing debt securities of the SPV from the trustor/seller.[184] In general, these securities were issued in tiers or tranches which were rated by various credit rating agencies, like Standard & Poor's, according to the likelihood that the holder would be paid (e.g., from "AAA" to "BBB" and below).[185] Before parting with higher ratings for senior investment grade securities, however, rating agencies would commonly require the SPV to obtain some form of credit enhancement, such as a guarantee, to ensure that payment shortfalls with respect to those securities would be covered.[186] For these senior tiers of debt, many SPVs looked to insurance companies to provide the requisite payment guarantee.[187]

---

180. *See* Taylor, *supra* note 15, at 878–82.

181. Kurt Eggert, *Held Up in Due Course: Predatory Lending, Securitization, and the Holder in Due Course Doctrine*, 35 CREIGHTON L. REV. 503, 571 (2002).

182. *See id.* at 572.

183. *Id.* at 538–39.

184. *Id.* at 539.

185. *Id.* at 540–41; Kathleen C. Engel & Patricia A. McCoy, *Turning a Blind Eye: Wall Street Finance of Predatory Lending*, 75 FORDHAM L. REV. 2039, 2047 (2007).

186. Engel & McCoy, *supra* note 185, at 2047.

187. *See* Christopher L. Peterson, *Predatory Structured Finance*, 28 CARDOZO L. REV. 2185, 2210 (2007).

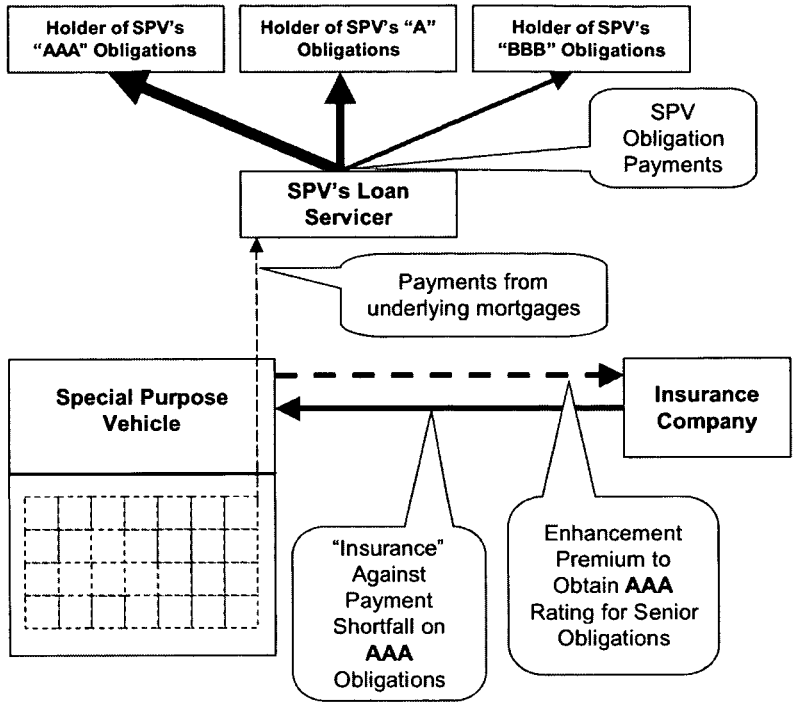


FIGURE 1: SPECIAL PURPOSE VEHICLES

At the same time, at the other end of the equation, financial institutions holding sub-prime, mortgage-backed debt obligations or some repackaged derivative thereof—such as a "collateralized debt obligation," or "CDO"[188]—would contract with counterparties to insure against default.[189] AIG Financial Products Corporation ("AIGFP"), a subsidiary of AIG, was one such counterparty.[190] AIGFP issued this form of "debt insurance"[191]—referred to as a "credit default swap," or "CDS"—even though it was organized as a bank and not an insurance company.[192] Though the CDS business generated hundreds of millions of dollars of insurance premium income when times were good,[193] the lurking obligations ultimately put AIG on the hook for various obligations under the CDOs.

188. Gretchen Morgenson, *Behind Biggest Insurer's Crisis, A Blind Eye to a Web of Risk*, N.Y. TIMES, Sept. 28, 2008, at A1, *available at* 2008 WLNR 18427841.

189. *Id.*

190. *Id.*

191. *See* Saporito, *supra* note 30, at 24 ("[A] CDS is, in its simplest form, an insurance policy . . . . AIG wrote multiple insurance policies covering the same underlying package of increasingly toxic assets.").

192. *Id.*; Henny Sender, *AIG Still Facing Huge Credit Losses*, FIN. TIMES, Mar. 4, 2009, at 1 ("AIG was particularly active in providing guarantees for . . . collateralised debt obligations, bonds backed by debts such as subprime mortgages.").

193. Morgenson, *supra* note 188, at A1 (indicating that the CDS business was generating as much as $250 million in premium income).

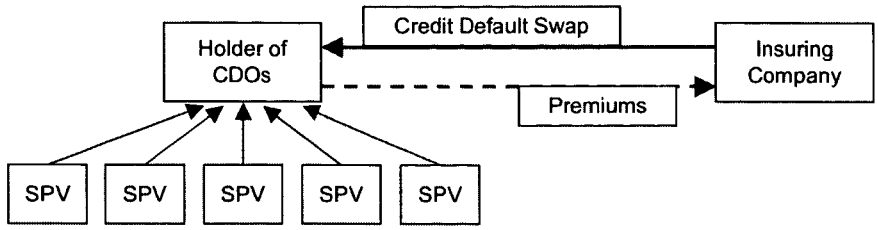


FIGURE 2: CREDIT DEFAULT SWAPS

With CDO stability ultimately depending on payment streams from holders of sub-prime mortgages and on rising home values, it did not take long for the disaster that was waiting in the wings to step in when the housing bubble burst.

From a pure insurance theory perspective, the problem is clear. Even with a large number of SPVs and an even larger pool of independent debt obligations insured against default, the law of large numbers can operate successfully only if the insurance issuer has reliable probability data with respect to the event insured against (i.e., default with respect to the obligation held).[194] Large numbers matter because such numbers generally make actuarial predictions meaningful.[195] Regarding pool size in this context, one commentator notes the following:

> [A]lthough credit derivatives may transfer insurance risk, an investor who assumes this risk does not usually assume the risks of a sufficiently large number of underlying obligors so as to pool and dilute (or diversify) the risks in any statistically meaningful sense. As a result, these investors do not satisfy the pooling condition for insurance and the contract should not constitute insurance for the credit protector or, under conventional wisdom, for the credit protected counterparty.[196]

The news gets worse from a probability data perspective. It might have been the case that historical mortgage default information was readily available and could have provided predictive value with respect to non-sub-prime mortgages. The same certainly cannot be said with respect to obligations backed by pooled sub-prime mortgages. Securitization of sub-prime mortgages is relatively new, and *a priori* probabilities do not exist with respect to the likelihood that this particular type of CDO will default. Acquiring a reliable default probability would have required an extensive assessment of the prior performance of similarly-backed CDOs, but at the time the "insurance" was issued, that information did not exist. According to at least one commentator, industry participants were, nonetheless, blindly undaunted:

194. *See* Taylor, *supra* note 15, at 877–83.

195. *See* Singer, *supra* note 15, at 116.

196. David S. Miller, *Distinguishing Risk: The Disparate Tax Treatment of Insurance and Financial Contracts in a Converging Marketplace*, 55 TAX LAW. 481, 542 (2002).

> Financiers, as well as the investors who bought their wares and the ratings agencies that evaluated them, agreed that by applying the proper equations it was possible to, say, bundle a bunch of subprime mortgages, chop them up and sell the pieces as fairly safe securities, even as they were leveraged to the hilt. Why? The mathematical models—backward-looking and based on just a few years' data from an asset bubble—said so.[197]

As it turns out, short-term experience proved negative for AIGFP, and, for the sake of the larger financial sector: AIG, the parent company ultimately liable, had to be rescued. To the extent that the prevailing "risk distribution" assessment is based on the mere existence of a large number of risk units in the insured pool, without any real effort to ascertain whether reliable risk-occurrence probabilities exist for that pool, the standard lacks real teeth and thus creates the appearance of insurance where none truly exists.[198] Indeed, investment activity can masquerade as "insurance." Ben Bernanke, Chairman of the Federal Reserve, hit the nail on the head when he referred to AIGFP as "a hedge fund basically that was attached to a large and stable insurance company."[199]

And that is not the worst of it from a federal tax perspective. Wholly aside from facts specific to AIGFP and AIG, if a given company's investment activities can mask as "insurance" (a likelihood enhanced by the proximity of legitimate insurance and the expectation of some investment activity), then a company may undeservedly qualify as an insurance company by appearing—to the naked eye—to satisfy the "more than half" standard. And even if a company satisfies the "more than half" standard, notwithstanding the masquerading of investment activity as insurance (because of the extent of its pure insurance business), the company nonetheless may understate its income by overstating its unearned premiums and/or unpaid loss reserves.

Although the Service tenaciously adheres to standards and concepts articulated in the 1940s to ascertain whether a given arrangement rises to the level of "insurance,"[200] modern courts have not hesitated to question

---

197. Barbara Kiviat, *Reassessing Risk*, TIME, Nov. 17, 2008, at 1, *available at* 2008 WLNR 21138582.

198. In discussing the concept of insurance for purposes of the McCarran–Ferguson Act, one commentator makes a similar point. *See* Steven J. Williams, Note, *Distinguishing "Insurance" from Investment Products under the McCarran–Ferguson Act: Crafting a Rule of Decision*, 98 COLUM. L. REV. 1996 (1998) (noting that the securitization process itself may appear to shift risk from one lender to another who pools and thereby distributes risk but does not, itself, constitute "insurance"). Highlighting the weakening of the line between banking and insurance, the author notes that "financial services innovation and expanding national bank powers have blurred the line between bank-authorized investment products and insurance." *Id.* at 1996. The author further points out that "a judicial definition of 'insurance' which only looks for some element of risk, without determining whether that risk is of the kind commonly associated 'insurance,' will inevitably reach beyond any common understanding of the term." *Id.* at 2019.

199. Sender, *supra* note 192, at 1.

200. *See* Rev. Rul. 2008-08, 2008-1 C.B. 340, 2008 WL 131946 (citing Helvering v. Le Gierse, 312 U.S. 531 (1941)).

the relevance of a risk-shifting inquiry.[201] In doing so, they rob that prong of the analysis of its evaluative force more generally. Further, even though there are those who embrace risk distribution as the true essence of insurance,[202] scholars compellingly argue that the law of large numbers has predictive value only when attached to reliable probability data,[203] a point which is occasionally lost on, ignored, or suppressed by modern-day insurers. Additionally, the law of large numbers does not appear to be a *sine qua non* of a risk distribution finding by the Service. It also bears noting that, in some contexts, adequate risk distribution is the only hurdle to clear. One commentator summarizes the problem well:

> One of the consequences of reliance on *Le Gierse* is that instead of developing analyses of pooling, homogeneity, independence and volume, advocates of both the economic family and recharacterization approaches are left with the bare words "risk distribution," which mean nothing without further explanation and contextualization. This problem was exacerbated in subsequent years when even the advocates of the recharacterization approach appeared to reduce the issue of defining an insurance operation to a computation of the amount of unrelated risks covered by the captive.[204]

Wholly aside from problems with the theoretical underpinnings of risk shifting and the limited utility of current risk distribution assessments, one can argue that both the two-test hurdle and the longstanding hostility towards deductions for self-reserving activity have resulted in an unhealthy concentration of risk (both predictable and dangerously unpredictable) in large, long-established entities that still qualify as "insurance companies," even though a hefty chunk of their business constitutes pure investment activity.

Shifting risk to an insurance company and having that risk distributed at the insurance company level may make good business sense for certain companies in certain industries. However, in viewing the universe of insurable risks more broadly, there is certainly room to consider approaches that promise better checks on risk concentration and thus, a more effective balancing of traditional risk shifting/distributing and rational risk retention.

IV. PROPOSALS

Scholars have proposed various solutions to the captive insurance company problem, such as focusing on the treatment of the captive and requiring the captive to satisfy certain requirements in order for it to

---

201. *See, e.g.*, Sears, Roebuck & Co. v. Comm'r, 972 F.2d 858, 863–64 (7th Cir. 1992).

202. *See* Note, *supra* note 19, at 784 (emphasizing that "[t]he process of risk distribution . . . is the very essence of insurance").

203. Taylor, *supra* note 15, at 878–83.

204. Singer, *supra* note 15, at 140.

qualify for special treatment—for example, proper risk-pooling, arms-length pricing, and maintenance of a sufficient level of unrelated policyholder business.[205] Fortunately, and perhaps due to a degree of judicial arm-twisting, some of those solutions are reflected in existing law. One approach, for example, respects the captive arrangement if there exists sufficient outside business at the captive level.[206] Others see the solution as a simple matter of weighing the business purpose of establishing the captive against any apparent tax avoidance purposes.[207] Neither approach is wholly satisfactory. The former fails to fully embrace the notion that one company may have a sufficient volume of homogenous risk and should have self-insuring options available without the participation of an external, unrelated third-party insurer. The latter approach is encouraging but could benefit from the imposition of specific standards to ensure that self-insuring is truly feasible, and not merely a matter of sufficiently weighty business purpose.

Thus, along those lines, the following discussion proposes contextual self-insuring. To address the problem of concentrated risk and to advance a simple utilitarian goal, this Article proposes the allowance of current deductions for the creation of limited contingency reserves for a specific subset of taxpayers.

A. Contextual Self-Insuring

The contextual self-insuring model seeks, at root, to allow companies with certain risks to self-insure through a related entity and deduct the relevant "premiums," even if the arrangement would not satisfy the Service's traditional risk shifting/distributing standard (such as a parent–subsidiary arrangement). A company would qualify for contextual self-insuring only if it could satisfy each of the following requirements:

(1) The arrangement involves an insurance risk and not merely an investment risk;

(2) Either the company finds a captive arrangement financially appealing *or* the desired insurance is commercially unavailable, available only at an exorbitant cost, or available at a cost substantially higher than the amount the company would have to pay if self-insuring;

---

205. *See, e.g., id.* at 162–164.

206. *See, e.g., Sears, Roebuck & Co.*, 972 F.2d at 863–64.

207. Winslow, *supra* note 15, at 135 (arguing that captive insurance arrangements should be analyzed by weighing legitimate business purposes against tax avoidance purposes, rather than attempting to ascertain whether insurance is "theoretically present"). Winslow has the same view with respect to retrospectively rated insurance contracts. *See* Donald Arthur Winslow, *A Note on Retrospectively Rated Insurance and Federal Income Taxation*, 79 KY. L.J. 195, 221–23 (1990) (suggesting that a business purpose test might be a more useful means of ascertaining whether insurance exists in a retrospectively rated contract). He notes that a retrospective insurance arrangement allows an insurance company to share the cash flow advantages typical of self-insurance or captive insurance company arrangements. *See id.* at 197.

(3) The taxpayer has a sufficient number of homogenous, independent risk units;

(4) There exists sufficient, reliable historical information regarding the risk pool such that the law of large numbers can operate effectively; and

(5) The insuring entity will be adequately capitalized and operated in accordance with the law of the captive's jurisdiction.

Companies unable to satisfy the foregoing would have the option of forming a group captive or going the traditional route and seeking coverage from an insurance company in an arrangement currently sanctioned by existing law. The proposal effectively allows individual companies that are able to distribute their own risks internally to avoid the high cost of obtaining that insurance from a third party, an alternative that is particularly appealing if the company is required to obtain the coverage. Further, this approach allows those same companies to access reinsurance markets directly, to the extent they feel reinsuring is appropriate, and avoids many potential problems that might be encountered when dealing with a third-party insurance company (such as resistance to claim payment and unnecessary delay).

Contextual self-insuring is, of course, vulnerable to the argument that allowing some companies to take deductions for self-insuring allows distortion and substantial understatement of income, a result that is unfair to other companies lacking the business profile or sheer size which would allow them to take advantage of this elaborate form of contingency reserving. There is also the potential charge that the eligibility standards are weighed down by ambiguities, such as: What qualifies as "exorbitant cost," when does a higher cost become "substantially higher," and what qualifies as "reliable historical information?"

Fortunately, current and historical market data can point those in a position to apply the standards in the right direction. Prevailing market rates can easily sort out "exorbitant" costs, and neutral decision-makers can readily assess the extent and quality of probability data to ascertain whether it is reliable and useful in the relevant context. That being said, a limited amount of ambiguity should not be fatal; applicable standards should allow some degree of flexibility and room for healthy evolution while, at the same time, retaining their immunity from easy manipulation. To the extent that there are legitimate fears regarding income distortion, there are a host of options with respect to reducing or minimizing the apparent harm. For example, allowing deduction of only a percentage of the premium, phasing in a full or partial deduction over a given time period, and disallowing a premium deduction altogether once the insuring company attains a certain financial profile.

While it remains true that only some companies will qualify for contextual self-insuring, society cannot deny the differential treatment of

various corporate taxpayers under current law, and should weigh the harm inadvertently suffered by individual companies against the overall health of the business sector and dangers of concentrated risk. Small companies, no matter how limited their product or service line, have insurable risks. Thus, to the extent these companies choose to grow and expand, they too may ultimately attain that critical mass of independent risk units which would allow them to take advantage of contextual self-insuring.

B. Utilitarian Contingency Reserving

The notion of utilitarian contingency reserving proposed here consciously departs from any consideration of risk shifting, risk distributing, or traditional notions of insurance. Rather, the core idea is that companies of a certain size—or those operating in certain industries—should, as a matter of sound tax policy, and in accord with pursuit of the common good, be allowed to establish and take deductions for contributions to a limited contingency reserve.

What industries and company sizes qualify? The question is tough to answer but obviously depends on the extent lawmakers wish to make available deductions for reserving. Whether broadly or narrowly tailored, if a utilitarian justification is to have force, a qualifying company must either be large enough such that a failure of the company or a sudden financial challenge to it would negatively affect a substantial segment of the U.S. economy (i.e., "too big to fail"), or be engaged in an industry in which the ready availability of capital in an emergency (or even over the long haul for crucial non-emergency activities) would maximize the health, safety, and welfare of the citizenry, such as clean-up immediately after an environmental disaster.

In any event, federally-endorsed reserving outside the insurance arena is not a revolutionary notion. Since 1984, taxpayers operating nuclear power plants have been able to take deductions for contributions to a Nuclear Decommissioning Reserve Fund ("NDRF" or "Fund"),[208] and the tax rules and procedures governing NDRFs serve as an excellent model for utilitarian reserving more generally.

Section §468A of the Code allows electing taxpayers to deduct amounts contributed to a Fund during the taxable year,[209] but limits the annual contribution to a designated "Ruling Amount."[210] Determined by

208. I.R.S. Chief Counsel Adv. 2007-03-007 (Jan. 19, 2007), 2007 WL 121781 (concluding that nuclear decommissioning cost is not an insurance risk and noting that "[t]he obligation to decommission has attached[,] therefore no hazard or fortuity as to the occurrence of decommissioning exists. . . . Thus, no insurance risk is involved.").

209. I.R.C. § 468A(a) (2006).

210. *Id.* § 468A(b).

an established schedule,[211] the Ruling Amount for each taxable year basically allows the taxpayer to contribute to the fund the total cost of decommissioning (but not more) over the estimated useful life of the power plant at a level-funding rate.[212] On the flip side, the electing taxpayer faces a gross income inclusion when amounts are distributed from the Fund[213] or, barring regulatory exception, certain other events occur.[214] Assuming the amounts were distributed from the Fund and used to cover actual decommissioning costs,[215] the gross income inclusion is offset by a deduction for decommissioning expenses when economic performance occurs.[216]

Although Congress designed the rules set forth in §468A for nuclear power plants facing definite decommissioning obligations, one could readily argue that a similar system would be useful for risks that are more contingent but financially radioactive, environmentally offensive, or threatening in their own right. If, for example, Congress allows companies to call on their contingency reserves (with an offsetting deduction for designated expenditures) as the economy *begins* to slip into a recession (or a given company encounters exigent financial circumstances), they may be able to delay or avoid layoffs, or take other steps to smooth out the volatility of the short- and long-term business cycles. And, to the extent that more large companies are able to bear their own risks, or at least fine tune the balance of risks shifted and those retained, American taxpayers will benefit as their shoulders are relieved of potential bailout burdens.

If the recent past is any indicator, federal bailout infusions, granted in exchange for an equity stake, exacerbate corporate agency costs. So, rather than those costs being borne by the historic shareholder group alone, those costs spread to the larger U.S. taxpaying population. It remains true that as with contextual self-insuring, utilitarian reserving can be attacked as allowing distortion and substantial understatement of income. However, the core premise of utilitarian thought is that the larger societal benefits outweigh individual harms.

---

211. *Id.* § 468A(d)(1). Although the Secretary is obligated to review the schedule of ruling amounts during the useful life of the power plant (revising them if necessary), the Secretary may do so more frequently if the taxpayer requests. *Id.* § 468A(d)(3).

212. *Id.* § 468A(d)(2)(A).

213. An exception applies with respect to amounts distributed from the Fund to cover certain costs connected with Fund operations. *Id.* § 468A(c)(1)(A), (e)(4)(B).

214. These events include deemed distributions under § 468A(e)(6), terminations of the Fund under § 468A(e)(7), and a disposition of any interest in the nuclear power plant. *Id.* § 468A(c)(1)(B).

215. The Fund can be used to cover decommissioning costs, to cover operational expenses, and to the extent of amounts not needed for the foregoing, to make investments. *Id.* § 468A(e)(4). For provisions governing the taxation of the Fund, see *id.* § 468A(e)(2).

216. *Id.* § 468A(e)(2)(A).

CONCLUSION

In light of general judicial rejection and scholarly criticism, the Service officially abandoned its economic family theory in the captive insurance company arena in 2001. As promised, however, the Service continues to analyze certain captive insurance transactions to ascertain whether they incorporate risk shifting and risk distributing, despite the fact that courts have questioned the legitimacy of a risk shifting inquiry and thereby weakened its determinative force. Commentators do not criticize the risk distribution requirement as inherently useless, but they do accurately note that the standard can be used to demarcate true "insurance" only in those contexts in which reliable probability data exists with respect to the designated risk pool.

To the extent the Service or any other entity focuses only on the mere existence of a large number of apparently independent, homogenous risk units in the designated pool (with very little or no regard for predictive data with respect to that pool), latent investment activity may successfully cloak itself in the garb of "insurance." AIG's so-called "debt insurance" on credit default swaps has brought home that truth and, given the concentration of risk within behemoth AIG, that truth arrived with resounding and devastating force.

Contextual self-insurance represents a healthy and viable means of de-concentrating insurable risk generally, and for companies deemed "too big to fail"—or those operating in certain industries—limited contingency reserving is far more palatable than multi-billion-dollar bailouts and the attendant enhancement of corporate agency costs. Indeed, both contextual self-insuring and limited contingency reserving promise to maximize the common good, whether measured on the national or individual community scale.

HE SPEAKS NOT, YET HE SAYS EVERYTHING; WHAT OF THAT?: TEXT, CONTEXT, AND PRETEXT IN *STATE V. JEFFREY DAHMER*

GREGORY J. O'MEARA, S.J.[†]

> [Dahmer] drill[ed] holes in his living victims' heads, pour[ed] in chemicals to "zombify" them, ha[d] sex with the corpses' viscera, and ke[pt] some body parts in his refrigerator, occasionally eating them.[1]
>
> Of course, in some respects, Abraham does speak. He says a lot. But even if he says everything, he need only keep silent on one single thing for it to be concluded that he hasn't spoken.[2]
>
> She speaks, yet she says nothing; what of that?[3]

INTRODUCTION

In *State of Wisconsin v. Dahmer*,[4] the defense attempted to lead the jury through a series of inferences that would have them conclude the defendant was insane at the time he committed each of the fifteen murders charged by the State of Wisconsin. They portrayed a client who cooperated fully with the authorities and who was too disturbed to be responsible for his actions. To make this approach work, they needed narrative distance between Dahmer and the jury so he would not be interrogated about his prior inconsistent statements and meticulous planning of the killings. This distance was created by Dahmer's silence in the courtroom. The jury heard his words only through others' voices. Though silence had worked as Dahmer's strategy outside the courtroom, the weight of the evidence undermined that approach at trial. His actions

---

† Assistant Professor, Marquette University Law School. B.A., Notre Dame; J.D., Wisconsin; L.L.M., New York University. When I was an Assistant District Attorney for Milwaukee County, I was part of the prosecution team in *State v. Dahmer*. Lead counsel for the prosecution was then Milwaukee County District Attorney E. Michael McCann, assisted by both Assistant District Attorney Carol L. White and me. The defense was headed by Mr. Gerald Boyle, assisted by Ms. Wendy Patrikus and Ms. Ellen Ryan. The current District Attorney of Milwaukee County, John Chisolm, graciously gave me access to the office's file in the *Dahmer* case. I am grateful to him and his assistant, Ms. Sheila Stanelle, for their generosity in the preparation of this article. I am grateful also to Professors Edward M. Gaffney, Bruce Berner, Scott Moss, Philip Chmielewski, S.J., Daniel Blinka, and Paul Secunda for their help on earlier drafts of this paper. Thanks also to Mr. Bryan Bayer and Mr. Michael Moeschberger for their research assistance.

1. Stephen M. Glynn, *If Dahmer's Not Crazy, Who Is?*, NAT'L L.J., Mar. 9, 1992, at 13.
2. JACQUES DERRIDA, THE GIFT OF DEATH AND LITERATURE IN SECRET 60 (David Wills trans., Univ. of Chi. Press 2d ed. 2008) (1999).
3. WILLIAM SHAKESPEARE, ROMEO AND JULIET act 2, sc. 2.
4. No. 1991CF912542 (Milwaukee County Cir. Ct. filed July 25, 1991).

spoke louder than his wordlessness, and the jury spoke in finding him responsible for what he did.

The job of a trial attorney is to tell a story and create a reality in the courtroom using the tools at hand: exhibits, testimony of witnesses, the rules of evidence, and the substantive law at issue in the case.[5] In crafting this narrative, attorneys take a complex set of events and filter them into various causal chains which are necessarily selective and stripped-down representations of what occurred on some prior date or series of dates.[6] Sensitive attorneys understand that success before a jury requires apprehension not only of content, the "what" of the narrative, but also of style, the "how" of the narrative.[7] This emphasis on style is critical because the incompleteness of information given to the jury requires it to fill gaps in reasoning.[8] Often attorneys selectively choose which facts are presented to a jury because there is too much material.[9] Such selectivity may also imply causal inferences in the jury.[10] This Article maintains that, in the case of *State of Wisconsin v. Jeffrey Dahmer*, how the case was presented was just as important as the content of that evidence. The text of the evidence needed a context; without it, the jury would not be persuaded.

In January 1992, television cameras and newspaper reporters flocked to Milwaukee, Wisconsin as the case pitting the State of Wiscon-

---

5. *See, e.g.*, ANTHONY G. AMSTERDAM & JEROME BRUNER, MINDING THE LAW 110–11 (2000); JAMES BOYD WHITE, WHEN WORDS LOSE THEIR MEANING: CONSTITUTIONS AND RECONSTITUTIONS OF LANGUAGE, CHARACTER, AND COMMUNITY 265 (1984) [hereinafter WHITE, CONSTITUTIONS AND RECONSTITUTIONS]; JAMES BOYD WHITE, *Rhetoric and Law: The Arts of Cultural and Communal Life*, *in* HERACLES' BOW: ESSAYS ON THE RHETORIC AND POETICS OF THE LAW 28, 34–35 (1985) [hereinafter WHITE, HERACLES' BOW].

6. *See* PAUL RICOEUR, MEMORY, HISTORY, FORGETTING 85 (Kathleen Blamey & David Pellauer trans., 2004) ("It is, more precisely, the selective function of the narrative that opens to manipulation the opportunity and the means of a clever strategy, consisting from the outset in a strategy of forgetting as much as in a strategy of remembering.").

7. *See, e.g.*, AMSTERDAM & BRUNER, *supra* note 5, at 167, 175 (discussing how rhetorical style can imply a range of meanings without saying any of them explicitly).

8. *See* SEYMOUR CHATMAN, STORY AND DISCOURSE: NARRATIVE STRUCTURE IN FICTION AND FILM 28–29 (Cornell Paperbacks 1980) (1978):

> Whether the narrative is experienced through a performance or through a text, the members of the audience must respond with an interpretation: they cannot avoid participating in the transaction. They must fill in gaps with essential or likely events, traits and objects which for various reasons have gone unmentioned. If in one sentence we are told that John got dressed and in the next that he rushed to an airport ticket counter, we surmise that in the interval occurred a number of artistically inessential yet logically necessary events: grabbing his suitcase, walking from the bedroom to the living room and out the front door, then to his car or to the bus or to a taxi, opening the door of the car, getting in, and so on. The audience's capacity to supply plausible details is virtually limitless, as is a geometer's to conceive of an infinity of fractional spaces between two points.

9. Indeed, trial courts demand that attorneys pare down facts to avoid repetition or waste of time. *See, e.g.*, Brown v. Wainwright, 785 F.2d 1457, 1466 (11th Cir. 1986) ("In the normal evidentiary sense cumulative evidence is excluded because it is repetitious."); *accord* Int'l Minerals & Res., S.A. v. Pappas, 96 F.3d 586, 596 (2d Cir. 1996).

10. *See* CHATMAN, *supra* note 8, at 28–29.

sin against serial killer Jeffrey Lionel Dahmer unfolded.[11] The defendant pled not guilty by reason of mental disease or defect to fifteen counts of first-degree intentional homicide.[12] Although he admitted killing the fifteen victims, he maintained he should not be held responsible for those deaths on the ground that he suffered from a mental disease, and, because of this disease, he was unable to conform his actions to the requirements of the law.[13] Essentially, Dahmer claimed he was a victim of his psychological disturbance and was no more to be blamed for his actions than were the young men whom he killed.

From one perspective, the trial should have been simple. Because of the guilty plea, there was no need for the panoply of witnesses and physical evidence that normally attends a homicide prosecution; no need for coroner reports to determine cause of death; no need for specific details of each of the fifteen murders because the defendant conceded causing them. Because the affirmative defense carries the civil burden, the defendant did not need to prove his case beyond a reasonable doubt[14] nor win the assent of a unanimous jury.[15] Rather, the jury had only to weigh the testimony of detectives, acquaintances of the defendant, and expert witnesses to determine the answers to two questions: at the time of each murder (1) "did the defendant have a mental disease or defect?" and, if that question were answered in the affirmative, (2) "[a]s a result of that mental disease or defect, did the defendant lack substantial capacity either to appreciate the wrongfulness of the conduct or to conform that conduct to the requirements of the law?"[16] Sentencing was not expected to be a major issue in this case—the defendant was going to be locked up for the rest of his life.[17] The sole question for sentencing was which sort of institution would house him: a prison or a hospital.[18]

---

11. Duane Dudek, *Dahmer's Insanity Defense Brings Court TV Coverage*, MILWAUKEE SENTINEL, Jan. 15, 1992, at 2 (describing how Court TV offered "gavel-to-gavel coverage" of the trial).

12. *Dahmer Changes Plea to Guilty but Insane*, N.Y. TIMES, Jan. 14, 1992, at A19, *available at* 1992 WLNR 3335725.

13. *Id.*

14. WIS. STAT. ANN. § 971.15(3) (West 2009) ("Mental disease or defect excluding responsibility is an affirmative defense which the defendant must establish to a reasonable certainty by the greater weight of the credible evidence."). The provisions of the Wisconsin Statutes dealing with the defense of mental disease or defect have not substantively changed since the Dahmer case was tried. *See, e.g.*, Act of May 27, 1994, 1993 Wis. Act 486, 1993 S.B. 826 (amendment to remove sexist language).

15. WIS. STAT. ANN. § 971.165(2) (West 2009) ("No verdict on the plea of not guilty by reason of mental disease or defect may be valid or received unless agreed to by at least five-sixths of the jurors.").

16. WIS. JI–CRIMINAL § 605, at 1 (2003). The jury's deliberation on this matter is directed by two questions on the verdict form. *Id.* at § 605B, at 1. The jury is directed "to answer the second question only if [they] answer the first question 'yes.'" *Id.* at § 605, at 1.

17. Pursuant to § 971.165, the jury was instructed that Jeffrey Dahmer would not go free if they found him not guilty by reason of mental disease or defect. Rather, he would be committed to the custody of the state and likely be confined to a state mental hospital "unless the court determines that the defendant would not pose a danger to himself or herself or to others if released under conditions ordered by the court." § 971.165(2). If the jury rejected the defense of not guilty by reason of

However, a close examination of the evidence presented in this case subverts any assertion of simplicity. In every insanity case, the legal and mental health professions understand mental disease differently; those difficulties were present here as well. Further, Dahmer's volubility and his penchant for documenting his actions with photographs and mementos required his attorneys to walk a fine line in crafting a picture of him that was at once familiar enough to garner jury sympathy and odd enough to assure a finding that he suffered from a mental disease. While proclaiming a strategy of complete and open disclosure, the defense carefully avoided facts that might have derailed the story it attempted to construct for the jury.[19]

This strategy of concealment is made manifest with a simple observation: Dahmer's guilty plea constituted the bulk of the words he said in open court before sentencing.[20] Although Dahmer's statements to detectives and mental health professionals provided almost all the facts assumed as true in the case, he was never sworn in as a witness, he never spoke at trial, and all of his words were mediated by others who reported them.[21] Rather than permitting the jury to observe the flesh-and-blood defendant from the witness stand, the defense orchestrated his previous statements to fabricate the most appealing figure possible, never undermining this discursive image with the defendant's live testimony. The strategy not only focused the jury but also controlled the defendant, whose earlier statements revealed inconsistencies and admissions detrimental to his defense.

After a brief description of the facts, and the substance of the testimony, this Article will trace the development of the insanity defense as it applies to the *Dahmer* case. I will then present challenges that confront any defendant who attempts to raise this argument in a criminal trial, coupled with challenges peculiar to defending Jeffrey Dahmer himself,

---

mental disease or defect, then the defendant would likely be sentenced to prison; specifically, in Wisconsin, he would be sentenced to life imprisonment. *See* WIS. STAT. ANN. § 939.50(3) (West 2009). Realistically, the issue at trial was not whether Jeffrey Dahmer would be locked up for life, but *where* this incarceration would take place.

18. *See* § 939.50(3); WIS. JI–CRIMINAL § 605, at 3.

19. This strategy of avoidance may also be attributed to the prosecution in this case. Although Mr. Dahmer confessed to a number of assaults in which he had sexual contact with people who were unconscious because he drugged them, the prosecution chose to focus solely on the murders in the complaint and information filed with the court. On the one hand, a number of murder charges surely outweigh incidental sexual crimes. On the other hand, it may be that the prosecution did not want to dilute the case's clean plot line by risking the jury's focus on the defendant's sexual desires and fantasies rather than on the tragic deaths of fifteen young men.

20. Obviously the defendant has a guaranteed right not to testify on his own behalf. Miranda v. Arizona, 384 U.S. 436, 442 (1966). Further, the prosecution cannot comment on the defendant's exercise of that right at trial. *See e.g.*, Doyle v. Ohio, 426 U.S. 610, 617–18 (1976). I do not challenge that right. Still, from the viewpoint of trial tactics and narrative theory, one cannot pretend that the defendant's choice not to take the stand is inconsequential. Part of what this Article does is show how far-reaching the decision to shift the defendant's words into other mouths at trial can be.

21. *See* Rick Romell & David Doege, *Boyle Says He Agonized over Dahmer Testifying*, MILWAUKEE SENTINEL, Feb 14, 1992, at 1.

given his actions before and after arrest. Next, I will turn to narrative theory to explain how the defense presented its case, demonstrating that its case-in-chief restricted the flow of information to the jury and painted a picture of the defendant at odds with his own statements. I will then sketch out how the prosecution countered this image.

I. THE VERDICT, AND THE FACTS UPON WHICH IT WAS BASED

The trial itself lasted fewer than thirteen days from opening statements to final summation. The jury's deliberation was complete in less than twenty-four hours. On February 15, 1992, Judge Laurence Gram, Jr. received a special verdict in which the jury found that, by a preponderance of the evidence, Jeffrey Dahmer did not suffer from a mental disease or defect when he committed the fifteen murders to which he had pled guilty.[22] The jury's finding was paradoxical because most of the psychiatrists and psychologists who examined Dahmer thought that he may have suffered from some sort of mental disease.[23] The verdict was met with cognitive dissonance both within the legal community and the public at large.[24] On the one hand, people appreciated that Jeffrey Dahmer was not "given a pass," and that he was held accountable for his disturbing actions which included murder, dismemberment, and cannibalism resulting in the deaths of at least seventeen young men in Ohio and Wisconsin between 1978 and 1991.[25] On the other hand, we feel uneasy with a finding that someone who kills seventeen people, collects the skulls and genitalia of some, eats parts of others, and drills holes in

---

22. Jim Stingl, *Dahmer Sane: Families of Victims to Describe Their Pain at Sentencing Monday*, MILWAUKEE J., Feb. 16, 1992, at 1. Because ten of the twelve jurors found that the defendant did not suffer from a mental disease, the jury did not consider the second question concerning the defendant's ability to conform his conduct to the requirements of the law. *Id.*; *see also* WIS. JI–CRIMINAL § 605, at 1.

23. As a matter of law, the jury instructions make clear:

> The term "mental disease or defect" identifies a legal standard that may not exactly match the medical terms used by mental health professionals. You are not bound by medical labels, definitions, or conclusions as to what is or is not a mental disease or defect to which the witnesses may have referred.

WIS. JI–CRIMINAL § 605, at 2. Still, even Dr. Park Elliott Dietz, one mental health witness who did not find that Jeffrey Dahmer suffered from a mental disease, wrote that a diagnosis of "a mixed personality disorder with antisocial, schizoid, and schizotypal features would be defensible." *See* Park Elliott Dietz, Report on Mental Status of Jeffrey L. Dahmer 6 (Jan. 10, 1992) (unpublished court file, on file with author).

24. My experience indicates that most people think Jeffrey Dahmer was successful in raising the defense of not guilty by reason of mental disease or defect. One may observe that the prosecution won where it counts—in court; but perhaps the defense did better than most thought in painting a picture of a troubled man, beset by a maelstrom of circumstances, rather than the somewhat cold and calculating killer that the prosecution argued was better supported by the evidence at trial. Despite winning the verdict, the prosecution's case seems not to have captured the popular imagination.

25. *See* Don J. DeBenedictis, *Sane Serial Killer: Experts Say Insanity Plea Alive and Well, Thanks Partly to Dahmer Jury*, A.B.A. J., April 1992, at 22. Theresa Smith, whose brother, Edward W. Smith was killed by Dahmer during the summer of 1990, noted that the verdict "brought back the faith I lost in the justice system." Dirk Johnson, *Milwaukee Jury Says Dahmer Was Sane*, N.Y. TIMES, Feb. 16, 1992, at 24, *available at* 1992 WLNR 3304512 (internal quotation marks omitted).

their skulls to "zombify" them is sane.[26] As one commentator asked, "If Dahmer's not crazy, who is?"[27]

Beginning with the death of Steven Tuomi in late 1987 and continuing until his arrest in July 1991, Jeffrey Dahmer refined his *modus operandi* for murdering men susceptible to his entreaties.[28] Dahmer would charm and seduce attractive young men, inviting them to come home with him and promising to pay them to pose for erotic photographs or to watch videos.[29] The victims were mostly in their twenties, and none of them drove a car.[30] After taking a cab or bus to a spot that was a few blocks from his residence,[31] Dahmer would walk to his dwelling with the victims, invite them in, and eventually offer them a drink laced with Halcion, a sleep aid for which Dahmer had a prescription.[32] Once the victims were unconscious, Dahmer would have sex with them, and then he would strangle them before they awoke.[33] Dahmer would often fondle their dead bodies and masturbate, and eventually he would move their bodies either to a drain spout or into a bathtub where he would cut them

---

26. Glynn, *supra* note 1, at 13.

27. *Id.*

28. Initially, Dahmer claimed all of his victims were gay or bisexual. Dennis Murphy, Statement of Jeffrey Dahmer, Case No. 2472 § 5, at 15 (July 23, 1991) (unpublished police report, on file with author). ("He stated the reason that he killed these homosexuals and he stated they all were homosexuals, was because he wanted to be with them."). Later, he admitted that a number of his victims were not gay but came home with him to be videotaped upon his promising them payment. Frederick A. Fosdal, Interview Notes for Jeffrey Dahmer Examination 23–25 (Nov. 13, 1991) (unpublished report, on file with author).

29. Dennis Murphy & Patrick Kennedy, Statement of Jeffrey Dahmer, Case No. 2472 § 5, at 88 (July 31, 1991) (unpublished police report, on file with author):

> As far as the sexual preference and/or race, religion, or education of the individuals that the suspect preferred, the suspect stated it was not a matter of race, religion, or education, it was just a matter of opportunity. He stated he offered each one of the individuals money to be photographed, to view videos, or to have sex, and after he persuaded them to come into his apartment, he would give them a sleeping potion, namely Halcion, and once they went to sleep, he would strangle them either manually or with a strap, photograph most of them after death, sometimes have sex with them after death, and then subsequently dismember them and on approximately eleven of the victims, kept the skulls and approximately four torsos, the hands, a couple hearts, and other inner organs.

30. Frederick A. Fosdal, Interview Notes for Jeffrey Dahmer Examination 59 (Jan. 10, 1992) (unpublished report, on file with author).

31. Murphy & Kennedy, *supra* note 29, at 81:

> He stated the reason why he would have the taxi drop him off several blocks from his apartment was in order to keep the taxi driver from knowing exactly where he lived at and to see if anyone had been following him, as he did not want anyone to detect his activities.

32. Dennis Murphy, Statement of Jeffrey Dahmer, Case No. 2472 § 5, at 149–50 (Aug. 22, 1991) (unpublished police report, on file with author):

> We then asked if he was experiencing any withdrawals, not only from alcohol but from not using halcion. He related that he has not experienced any withdrawals from alcohol nor from the use of halcion, because he does not take the pills regularly. He related that he would take one pill about every six months and that was only when he could not sleep. He related that the main reason he had halcion was to use [it] on the people he brought to his apartment or the ones he met in the bathhouses. He related that he first started to experiment by using three pills on the people and then used as many as seven on some of them. He related that he would bluff the doctors into prescribing the pills for him because he would tell them that he could not sleep but never used them.

33. Murphy & Kennedy, *supra* note 29, at 88.

up to dispose of them, occasionally saving trophies such as their skulls or preserved genitalia.[34] On four occasions he engaged in cannibalism but later stated he found this unfulfilling.[35] He would either burn or throw out their clothing and destroy any identification they had on them.[36] When Dahmer was arrested, there were remains of eleven of his victims in his apartment.[37]

In addition to these facts recounted by the defendant to police detectives, additional claims emerged from Dahmer's discussions with clinicians. Dahmer reported he attempted to exhume a freshly dead corpse for sexual purposes,[38] he drank blood from a test tube while working as a phlebotomist,[39] and he drilled small holes into the skulls of five of his victims while they were drugged and injected a mixture of muriatic acid and water, or boiling water alone, in an attempt to make them sexual slaves.[40] He also claimed that he planned to build a "'temple' that fea-

---

34. Frederick A. Fosdal, Interview Notes for Jeffrey Dahmer Examination 37 (Dec. 20, 1991); Patrick Kennedy, Statement of Jeffrey Dahmer, Case No. 2472 § 5, at 123 (July 23, 1991) (unpublished police report, on file with author).

35. Murphy, *supra* note 32, at 151–52:

> Jeff Dahmer went on to relate that he had originally told us that he had only eaten a bicep of one of his victims. He related that there were other times in which he had eaten part of the victim. The first time was the person he identified as Cash D (Raymond Smith—Victim #5). He related that he eat [sic] this victim's heart. He related that it tasted kind of spongy. He indicated that the next victim was the person he met by the bookstore, (Victim #7—Ernest Miller). He related that this was a person he really liked. He indicated that he had filleted his heart and had kept it in the freezer and also kept his bicep. He indicated that he had eaten the thigh muscle of this subject, but it was so tough he could hardly chew it. He then purchased a meat tenderizer and used it on the bicep. He stated that it tasted like beef or a filet mignon. The next person he was going to eat, and in fact tried, was victim # 15—Oliver Lacy. He stated that on this victim he ate his bicep. This also tasted like Filet mignon. He stated that he would tenderize it first. He stated that he did keep this individual's heart and bicep. We asked him if he had eaten the body parts, just plain. He stated that he would use salt, pepper, and A–1 Steak sauce on them. He stated that the reason he ate these parts was because he was curious but then it was because he wanted to make them a part of him. He stated that this way he could keep these people with him. He stated that he ate only the people that he really liked and wanted them to be a part of him or with him all the time.

See also Dietz, *supra* note 23, at 4 ("He had no enduring interest in cannibalism, but rather tried it out of curiosity and made use of the occasion to masturbate to fantasies of a victim he had consumed. Although he did so on as many as 10 occasions, this did not develop into an enduring or intense sexual interest.").

36. R.W. Munsey, Statement of Jeffrey Dahmer, Case No. 2472 § 12, at 195 (July 26, 1991) (unpublished police report, on file with author) ("Jeffrey L. Dahmer stated he took the victim's clothing to a location used to burn trash by his father where there he burned the clothing and identification.") (describing Dahmer's destruction of evidence following the Hicks homicide in Bath, Ohio).

37. *See* Jeffrey Jentzen et al., *Destructive Hostility: The Jeffrey Dahmer Case: A Psychiatric and Forensic Study of a Serial Killer*, 15 AM. J. FORENSIC MED. PATHOLOGY 283, 292 (1994).

38. Dietz, *supra* note 23, at 3.

39. *Id.* at 4.

40. Fosdal, *supra* note 28, at 26–27:

> "I was trying to think of a way to not have to kill them." . . . Said he would drill a small hole through the top of the skull and into the brain—on about four or five victims—used a baster . . . [and] injected diluted muriatic acid into the skull. . . . Said drilling the hole in the skull was "an experiment—it never worked out." Tried this technique on his last 4–5 victims—put acid into the hole on four of the victims and boiling water in one of the vic-

tured his victims' remains in hopes of 'receiving special powers and energies.'"[41]

II. THE UNDERLYING ANTHROPOLOGY OF THE CRIMINAL LAW

The above facts surely signal someone who is seriously disturbed. Indeed, if one asked someone on the street whether a person who did these things were crazy, the answer would be a resounding "yes."[42] This recognition gains significance in light of two seemingly contradictory positions held by the law. On the one hand, the criminal law prides itself on being a system that concerns itself with justice rather than vengeance.[43] Judgments are not determined by categories used by persons in ordinary discourse; the law calls for analytical distinctions developed throughout the course of its history.[44] On the other hand, juries determine if a defendant raises the insanity defense successfully, and the courts' instructions advise the panel that it may disregard expert witnesses' opinions and draw its own conclusion as to the mental state of the defendant.[45] These instructions invite jury members to give weight to their own reasoning and conclusions, even when those conclusions deviate from the conclusions of those recognized as experts in the field of mental health.[46] Because it grants jurors the authority to disregard expert opinions, the court must control precisely what information it permits the jury to consider. The insanity defense rests upon the Aristotelian assumption about all human beings that underlies the criminal justice system; specifically, the law presumes that human beings are rational and make free and unconstrained choices in this world.[47] Aristotle maintained that

---

tims. Said the five individuals did not get into a "zombie" state. . . . The purpose was that they would be alive, the bodies would be preserved, but their personality would be "zombie—so I wouldn't have to go out looking for partners."

41. David Doege, *Dahmer Planned Shrine of Bones*, MILWAUKEE SENTINEL, Feb. 5, 1992, at 1.

42. *See* GEORGES CANGUILHEM, THE NORMAL AND THE PATHOLOGICAL 117 (Carolyn R. Fawcett trans., Zone Books 1989) (1966) ("When we call another man insane, we do so intuitively 'as men, not as specialists.' The madman is 'out of his mind' not so much in relation to other men as to life: he is not so much deviant as different.").

43. *See* PAUL RICOEUR, CRITIQUE AND CONVICTION 117 (Lawrence D. Kritzman ed., Kathleen Blamey trans., Columbia Univ. Press 1998) (1995) ("[J]ustice encounters its contrary first in the thirst for vengeance, which is a powerful passion: justice consists in *not* seeking vengeance. Between the crime and the punishment, to return to well-known categories, lies justice and, consequently, the introduction of a third party.").

44. *See* RONALD DWORKIN, LAW'S EMPIRE 99 (1986) (discussing precedent and legislation in the conception of state power).

45. WIS. JI–CRIMINAL § 605, at 2 (2003) ("You are not bound by medical labels, definitions, or conclusions as to what is or is not a mental disease . . . ."). In an explanatory footnote to this instruction, the Jury Instructions Committee explains:

> The intent of this sentence is to emphasize that the jury is not bound by what is considered "mental disease or defect" for medical purposes. The jury *is* bound by the legal definition of mental disease as explained in this instruction. In a proper case, the judge may wish to emphasize this distinction.

Id. at 7 n.7.

46. *See id.* at 2, 7 n.7.

47. *See* ARISTOTLE, NICOMACHEAN ETHICS 65 (Martin Ostwald trans., Bobbs–Merrill Educ. Publ'g 1962):

the origin of our actions is internal, within ourselves, and thus voluntary: "[I]f we cannot trace back our actions to starting points other than those within ourselves, then all actions in which the initiative lies in ourselves are in our power and are voluntary actions."[48] Professor George Fletcher shows how this notion of the rational human being acting voluntarily is often formulated in terms of "free will."[49] Criminal law presumes free will operates in all situations unless we recognize extenuating circumstances giving rise to an excuse or justification.[50]

Despite the admonition that the law is addressed solely to rational actors choosing freely among alternative courses of action, the law has not always recognized an actor's rationality.[51] Rather, the law's early focus was solely on consequences of the physical act itself. In early English law, if there were a quarrel and a dead body resulted, then the killer, regardless of reasons for doing so, was liable to punishment.[52] As common law was influenced by canon (church) law following the Norman Conquest, matters began to change; canonists assigned weight not only to the act itself but also to the intention that lay behind it.[53] By determining a penitent's intention, the confessor could assign an appropriate pen-

---

> For where it is in our power to act, it is also in our power not to act, and where we can say "no," we can also say "yes." Therefore, if we have the power to act where it is noble to act, we also have the power not to act where not to act is base . . . . But if we have the power to act nobly or basely, and likewise the power not to act, and if such action or inaction constitutes our being good and evil, we must conclude that it depends on us whether we are decent or worthless individuals.

See also GEORGE P. FLETCHER, THE GRAMMAR OF CRIMINAL LAW: AMERICAN, COMPARATIVE, AND INTERNATIONAL 9–10 (2007) (discussing the history of Aristotle's view of involuntary actions by reason of mistake, duress, or insanity).

48. ARISTOTLE, *supra* note 47, at 65.

49. FLETCHER, *supra* note 47, at 10 ("This problem of attributing agency has traditionally been addressed under the label of 'free will.' In the Christian West, the discussion of free will took the place of Aristotle's focus on the issue of voluntary action.").

50. *See id.*

51. *See, e.g.*, GEORGE P. FLETCHER, A CRIME OF SELF DEFENSE: BERNHARD GOETZ AND THE LAW ON TRIAL 30 (1988):

> From roughly the 13th to the 16th century, the plea of self-defense, called *se defendendo*, came into consideration whenever a fight broke out and one party retreated as far as he could before resorting to defensive force. His back had to be literally against the wall.
>
> If he then killed the aggressor, *se defendendo* had the effect of saving the defendant from execution, but it left intact the other stigmatizing effects of the criminal law. The defendant forfeited his goods as expiation for having taken human life. The murder weapon was also forfeited to the Crown as a deodand, a tainted object. Killing *se defendendo* was called excusable homicide, for though the wrong of homicide had occurred, the circumstances generated a personal excuse that saved the manslayer from execution.

52. *See id.*

53. JOHN MAHONEY, THE MAKING OF MORAL THEOLOGY: A STUDY OF THE ROMAN CATHOLIC TRADITION 180 (1987). At least Aquinas considered both the act and the intention in determining the moral goodness of an action. His perspective was a corrective on the work of Peter Abelard which focused solely on the intention with which an act was done. Mahoney notes:

> [I]n his ethics Abelard was equally individualistic, to the extent of concentrating the morality of good or bad action not in what was being done, but in the intention with which it was done. . . . Moral goodness or badness does not reside in any action considered in itself but derives only from the intention which produces the action.

Id. at 176.

ance in the confessional.[54] Professor Fletcher maintains that the criminal law was influenced by these pastoral attempts at grading the intention behind an act.[55] For example, distinctions in self-defense can be traced back to distinctions made by canonists[56]:

> The modern approach to the distinction between self-defense and punishment finds its best exposition in the work of Thomas Aquinas, who emphasizes the intention with which the defender harms the aggressor. If the intention is not to harm but merely to fend off the attack, then the action can properly be described as an act of self-defense; but if the intention is to make the aggressor suffer for his misdeed, then the act appears to be closer to punishment. . . . The basic principle is that a private individual may not intentionally kill another human being when the explicit object (rather than the side effect) of the action is to cause death.[57]

Because the criminal law lacks ministers to interrogate the offender with the breadth and depth granted to the confessor,[58] it modified the canonist's approach to intent or motive. Rather than focusing on a psychological reality underlying a given act, criminal law constructs an element called the actor's "intent." The best definition of intent in criminal law arises in the work of the nineteenth-century jurist James Fitzjames Stephen: "The only possible way of discovering a man's intention is by looking at what he actually did, . . . what must have appeared to him at the time the natural consequence of his conduct."[59] The understanding of a defendant's intent is not an inquiry into the actual motives that stirred the defendant to action or deep desires of the heart; rather, it is an act of reconstruction based on the defendant's external actions.[60] The law looks at what an actor did and reasons backwards, presuming that the rational actor intended the ensuing consequences.[61]

By considering the actor's intent, the law went beyond its earlier consideration of external acts and focused on the actor's point of view. This concentration on the subject's motivation gave rise to two broad categories of defenses: justifications and excuses.[62] These defenses in

---

54. *See id.* at 179–80.

55. FLETCHER, *supra* note 47, at 14. ("[W]hat one intends specifies moral actions, not what one does not intend, since the latter result is accidental. . . . And so such acts of self-defense by them to preserve one's life, do not have the character of being unlawful" (alteration in original) (internal quotation marks omitted) (quoting St. Thomas Aquinas's *Summa Theologica*)).

56. *See id.*

57. *Id.* (footnote omitted).

58. *See* MICHEL FOUCAULT, THE HISTORY OF SEXUALITY 58–73 (Robert Hurley trans., Pantheon Books 1978) (1976).

59. *See* 2 JAMES FITZJAMES STEPHEN, A HISTORY OF THE CRIMINAL LAW OF ENGLAND 111 (1883).

60. *See id.*

61. *See id.*

62. Of course other defenses are possible, for example, those rooted in a failure to fulfill the elements of the crime or claims that the crime, as written, violates constitutional protections. Be-

turn imply the presence of an event conveyed to the jury by means of narrative.[63] They provide a context, a world that is subject to limitations, in which the defendant lives and in which he makes rational choices.[64]

III. JUSTIFICATIONS AND EXCUSES IN CRIMINAL DEFENSE

Justifications take the following form: "When viewed in its entirety, the defendant's act was neither wrong nor bad; indeed, the act was virtuous."[65] An act such as the use of defensive force to repel an aggressor's unlawful attack is an example of a justified action. We say that the actor is justified because we find self-defense or defense of others understandable, rational, and worthy of commendation.[66]

Excuse in criminal law is different. The law excuses criminal defendants from penal consequences for wrongful acts arising either through no fault of their own or in situations where the law perceives that the defendant was subject to a "maelstrom of circumstance."[67] Excuse is rooted in the sense that "what the actor did was wrong, but she had a good reason for doing it."[68]

Excuses in criminal law stem from reasons either external or internal to the actor. Duress is one example of an external force resulting in an excuse.[69] Assume defendant D shoots an innocent victim V, causing bodily injury to her. Normally this action should result in a charge of

---

cause this paper considers issues of defect in a defendant's intent, I am focusing on justification and excuse because these defenses are addressed particularly to the *mens rea* element.

63. *See, e.g.*, Robert M. Cover, *Nomos and Narrative*, 97 HARV. L. REV. 4, 4–5 (1983):

> No set of legal institutions or prescriptions exists apart from the narratives that locate it and give it meaning. For every constitution there is an epic, for each decalogue a scripture. Once understood in the context of the narratives that give it meaning, law becomes not merely a system of rules to be observed, but a world in which we live.

(footnote omitted).

64. *Id.*

65. GEORGE P. FLETCHER, RETHINKING CRIMINAL LAW 759 (1978):

> Claims of justification concede that the definition of the offense is satisfied, but challenge whether the act is wrongful; claims of excuse concede that the act is wrongful, but seek to avoid the attribution of the act to the actor. A justification speaks to the rightness of the act; an excuse, to whether the actor is accountable for a concededly wrongful act.

66. To label a defense as a justification rather than an excuse is already to accept a set of values subject to the vagaries of history. As alluded to above, the common law did not view as self-evident that killing another in self-defense freed the actor from punishment. *See* ARISTOTLE, *supra* note 47, at 65. The underlying command "Thou shalt not kill" would have been seen as outweighing an assumption that self-preservation is an unmitigated good. Only when moral weight is given to preservation of life as a good in itself can the act of defensive force be seen as justified. It may be that the intuitive "rightness" that characterizes a particular defense as a justification rather than an excuse is as much a product of social mores as anything else.

67. FLETCHER, *supra* note 65, at 808:

> Excuses are motivated by compassion for persons caught in a maelstrom of circumstance. The underlying sentiment is that if any one of us were forced to act at gunpoint or to steal in order to survive, we would do the same. If we recognize our essential equality with the accused and identify with his situation, then we cannot help but feel compassion and excuse his all-too-human transgression.

68. *See id.* at 802–03.

69. *See, e.g., id.* at 829–33.

battery. However, when D pulls the trigger and shoots V solely because actor A has a revolver cocked and pointed at D's temple, threatening to kill him if he fails to fire the gun, the law may excuse D's action. A's training a gun on D impeded D's ability to choose, and therefore, shooting the gun at V was not the product of D's will. To put the matter differently, D does not evince a criminal character by acting as he does, and his act may therefore be excused.

Excuses also arise from internal forces understood as burdening an actor's freedom of choice just as much as a gun held to his temple. Consider the excuse rationale underlying the mistaken use of defensive force. Imagine defendant E reasonably believes that her life is threatened by actor F. Assume further that E believes her failure to take immediate action to thwart this deadly attack will result in her death. Assume finally that E's belief is mistaken; she is not under attack at all. E's act of violence directed at F injures an innocent person. In this situation, the law may excuse E's action because, had the circumstances been as E reasonably believed them to be at the time of her act, E's actions would have been justified.

Excuse defenses uphold the anthropology of the rational actor because actors who are excused when under duress or mistaken about surrounding facts still act rationally.[70] They choose among alternatives after weighing options they perceive, even if their assumptions are later disproved. The plea of not guilty by reason of mental disease or defect likewise upholds the anthropology of the rational actor in criminal law because it excuses defendants who are unable to act rationally in a given situation through no fault of their own.[71] That said, the insanity defense remains filled with difficulties and seeming contradictions.[72]

IV. A BRIEF HISTORY OF THE CRIMINAL EXCUSE OF INSANITY

A. Vicious Wills and Reason

The first English lawyer to consider the mental element of the crime and propose relief for the insane was Henri de Bracton in his thirteenth century treatise *On the Law and Customs of England.*[73] He maintained

---

70. *Id.* at 802–03.

71. *Id.* at 835:

> The definition, administration and ramifications of the insanity defense express the deepest concerns of the Anglo-American legal culture. . . . In posing the question whether a particular person is responsible for a criminal act, we are forced to resolve our doubts about whether anyone is ever responsible for criminal conduct. And if some are responsible and some are not, how do we distinguish between them?

72. *See id.*

73. *De Legibus et Consuetudinibus Angliae* was completed around 1256; it is credited as "the first systematic treatise on English." Anthony Michael Platt & Bernard L. Diamond, *The Origins and Development of the "Wild Beast" Concept of Mental Illness and Its Relation to Theories of Criminal Responsibility*, 1 J. HIST. BEHAV. SCI. 355, 356 (1965); *see also* JOEL PETER EIGEN, WITNESSING INSANITY: MADNESS AND MAD-DOCTORS IN THE ENGLISH COURT 35 (1995):

that one needs a "will to harm" before a crime can be committed.[74] Drawing on canon law, Bracton noted that, just as the law does not hold infants or brute beasts responsible for the consequences of their behavior, there are some adults who should likewise be excused because their ability to reason is impaired, and they are thus comparable to children or "brutes."[75]

The focus on the will noted by Bracton was underscored later by Blackstone: "So that to constitute a crime against human laws, there must be, first, a vicious will; and, secondly, an unlawful act consequent upon such vicious will."[76] Although he agreed with Bracton on the necessity of the will in making legal determinations, Blackstone reframed the issue in terms of cognitive impairment.[77] Whereas Bracton excuses the insane whom lack "corrupt intent," "will to harm," and "malice," Blackstone addressed the actor's inability to reason because, presumably, reason informs the will:

> [I]f there be any doubt, whether the party be *compos* or not, this shall be tried by a jury. And if he be so found, a total idiocy, or absolute insanity, excuses from the guilt, and of course from the punishment, of any criminal action committed under such deprivation of the senses; but, if a lunatic hath lucid intervals of understanding, he shall answer for what he does in those intervals, as if he had no deficiency.[78]

According to Blackstone, a defendant may be relieved of responsibility not only when he manifests a wholly deficient reason but also at sporadic points when his reason seems impaired.[79] Further, during times when his reason seems unaffected, a defendant should be held responsible for what he does.[80]

B. Reason and the M'Naghten *Standard*

This recognition that partial impairment may support a defense of insanity occurs in the case of Daniel M'Naghten.[81] Here the test for insanity shifts from the defendant's volitional impairment to his cognitive

---

> Bracton . . . was the first English lawyer (he was also chancellor of Exeter Cathedral *and* chief justiciary of the highest court in the realm) to incorporate the mental element into legal writing: "For a crime is not committed unless the will to harm be present . . . In misdeeds we look to the will and not the outcome." In essence, the law conceived of people as capable of free choice, a free exercise of the will.

(alteration in original).

74. EIGEN, *supra* note 73, at 35.

75. *Id.* (internal quotation marks omitted).

76. WILLIAM BLACKSTONE, 4 COMMENTARIES *21.

77. *See id.* at *25.

78. *Id.*

79. *Id.*

80. *Id.*

81. M'Naghten's Case, 8 Eng. Rep. 718 (1843). *M'Naghten* was a politically celebrated case. *See* EIGEN, *supra* note 73, at 153.

processes. M'Naghten shot Edward Drummond who eventually died from the wound.[82] M'Naghten was charged with murder and pleaded not guilty by reason of insanity.[83] The case presented difficulties because M'Naghten fit into neither of the two categories described by Blackstone. He was neither a "lunatic" nor someone who seemed insane with occasional lucid intervals.[84] Rather, M'Naghten appeared otherwise sane and lucid except when dwelling upon a particular delusion that he was the victim of political persecution.[85] The physician Edward Thomas Munro testified in the trial court that this condition should be sufficient to relieve the defendant of responsibility:

> [A] person may have a morbid delusion, and yet still know that thieving is a crime, or that murder is a crime, but his antecedent delusions lead to one particular offense or another . . . [I] think that delusion of this nature [political persecution] carries a man quite away—I mean that his mind was so absorbed in the contemplation of the fancied persecution, that he did not distinguish between right and wrong.[86]

In its recitation of the facts, the high court agreed as it observed:

> [I]t was of the nature of the disease with which the prisoner was affected, to go on gradually until it had reached a climax, when it burst forth with irresistible intensity: that a man might go on for years quietly, though at the same time under its influence, but would all at once break out into the most extravagant and violent paroxysms.[87]

M'Naghten's case perplexed the Queen's Bench because his customary appearance of sanity raised the possibility of the defendant's lying.[88] The original jury returned a verdict of not guilty by reason of in-

---

82. *M'Naghten*, 8 Eng. Rep. at 718. Drummond was the prime minister's secretary; most people think M'Naghten mistook Drummond for Prime Minister Peel, a fact which cuts both for and against his sanity. *Id. See* ROGER SMITH, TRIAL BY MEDICINE: INSANITY AND RESPONSIBILITY IN VICTORIAN TRIALS 124–25 (1981) (discussing the development of the *M'Naghten* rules in light of competing professional claims by medicine and law).

83. *M'Naghten*, 8 Eng. Rep. at 718.

84. *See* BLACKSTONE, *supra* note 76, at *25.

85. *See* EIGEN, *supra* note 73, at 153–54 (citing Old Bailey Sessions Papers, Case 874, 5th Sess. 756–59 (1842-43).

86. *Id.* at 153 (alteration in original) (internal quotation marks omitted) (quoting Old Bailey Sessions Papers, Case 874, 5th Sess. 756–59 (1842–43)).

87. *M'Naghten*, 8 Eng. Rep. at 718.

88. In recognizing this possibility, the court touched upon a widely-held belief that still infects the public perception of the criminal justice system. *See, e.g.*, CHRISTOPHER CRONIN, FORENSIC PSYCHOLOGY 93 (2006):

> When asked, most students will estimate that the insanity plea is used anywhere from 25 to 50 percent of the time in criminal cases. Additionally, the public also feels that it is generally successful as a way to avoid incarceration. One study found that the public thought that the insanity plea was used as a ploy in nearly 50 percent of all criminal cases and that it was successful 20 percent of the time. Actually, the plea of not guilty by reason of insanity is used in less than 1 percent to 3% of all criminal cases. . . . It also has a much lower success rate than most people believe. Several studies have found that the plea is successful . . . approximately 25 percent of the time. . . . Approximately 70 percent of the insanity acquittals are the result of a plea bargain or similar arrangements rather

sanity.[89] This decision was appealed to the House of Lords, which, after debating the matter, referred the case for declaratory judgment to the High Court.[90]

The published decision consists of two opinions.[91] In the initial opinion, Mr. Justice Maule ruled that the insanity defense is available when there is proof of the unsoundness of mind "such as render[s] [a defendant] incapable of knowing right from wrong."[92] Lord Chief Justice Tindal, writing for the majority, took Maule's idea and developed it at greater length. Tindal noted that the law presumed every defendant's sanity; therefore, the defense had the burden of proving to the jury's satisfaction that a party was "labouring under such a defect of reason, from disease of the mind, as not to know the nature and quality of the act he was doing; or, if he did know it, that he did not, know he was doing what was wrong."[93]

The rule makes two distinctions. In the first instance, one fails "to know the nature and quality of the act he was doing" insofar as the actor is so deluded that he truly believes he is performing one action when he is doing something wholly other.[94] Tindal also proposed a separate category of defendants who should not be held responsible: those who knew what they were doing but did not know that it was wrong.[95] Furthermore, Tindal indicated that the mere presence of a delusion in the mind of the defendant is insufficient to grant the defense.[96] Rather, Tindal directed trial courts to consider the nature of the delusion and how it affects the actions of the defendant.[97] A defendant may only be excused if, were his delusion correct, he would have had an excuse under the law, such as that granted by a reasonable mistake of fact.[98]

The development of the law from Blackstone to *M'Naghten* parallels the emergence of psychiatry as a profession.[99] Psychiatric experts changed the quality of evidence sufficient for excusing criminal liability on the grounds of insanity.[100] Previously, the finding of insanity rested

---

than through a jury trial. This is not too surprising in light of the fact that juries tend to hold negative attitudes toward the insanity defense.

(citations omitted) (emphasis omitted).

89. *M'Naghten*, 8 Eng. Rep. at 718–19.

90. *Id.* at 719.

91. *See id.* at 719, 721.

92. *Id.* at 720.

93. *Id.* at 721.

94. *Id. See generally* OLIVER SACKS, THE MAN WHO MISTOOK HIS WIFE FOR A HAT (Touchstone 1998) (1985) (describing victims of neurological disorders so profound that they simply cannot distinguish among various objects and actions).

95. *See M'Naghten*, 8 Eng. Rep. at 721.

96. *See id.* at 721.

97. *See id.* at 722.

98. *Id.*

99. *See e.g.*, SMITH, *supra* note 82, at 90–91.

100. *See* ALAN NORRIE, CRIME, REASON AND HISTORY: A CRITICAL INTRODUCTION TO CRIMINAL LAW 173 (Robert Stevens et al. eds., 1993):

upon the ordinary observations of lay people.[101] After the rise of psychiatrists (or alienists, as they were called) in the nineteenth century, the finding of insanity became problematic; mental illness was no longer something seen by ordinary people. Rather, mental illness referred to something more occult, internal, and observable only by those with special training.[102]

C. Reason, Mental Illness, and the Products Test

The *M'Naghten* rule seized the legal imagination and became the test to determine criminal insanity not only in England but also in the vast majority of United States jurisdictions. In 1972, the American Law Institute's Model Penal Code revised and added to the *M'Naghten* test, which resulted in "the ALI test".[103] The ALI test continues the cognitive element of the *M'Naghten* rule[104] and adds a volitional element, which

---

> The concept of insanity appears to fit neatly into an orthodox liberal framework. The insane person is morally, therefore legally, irresponsible for his acts, and unpunishable. At most, the criticism might be that the law's outmoded narrowness stems from a judicial over-sensitivity to the needs of social protection which should be corrected by reform in favour of the accused. But what cannot be recognised from this perspective is how the traditional views about insanity are *ideologically* entrenched within legal discourse, so that much more rides on the issue than a small measure of enlightened liberal reform. At stake is a particular way of seeing the social world and the human beings that populate it that is both powerful, and odd.

(citation omitted).

101. The ordinary observations of lay people still constitute valid evidence of a defendant's responsibility when he pleads not guilty by reason of insanity. *See* Duthey v. State, 111 N.W. 222, 225–26 (Wis. 1907). Relying on *Duthey*, the prosecution in the *Dahmer* case repeatedly asked ordinary lay people about their observations of Jeffrey Dahmer and if they thought he was mentally ill on the basis of their experience.

102. *See, e.g.*, NORRIE, *supra* note 100, at 176:

> Insanity came increasingly to be seen as a product of disease located in the brain which caused the mad behaviour. Following their methodology to its natural conclusion, psychiatrists then argued that the 'truth' of insanity lay not in its empirical manifestation, in conduct displaying an obvious lack of reason, but in the underlying causal mechanisms to be found in the brain.
>
> It was this move in thinking that caused the break with law. If the ultimate locus of insanity was not in its psychological manifestation but in underlying organic causes, it became possible to conceive of forms of insanity which left the 'surface' areas of the psyche, for example the reasoning faculty, relatively unaffected while attacking the 'deeper' elements of the will or the emotions. A lack of reason became one, but only one, *symptom* of an underlying, causal, mental illness. A man could as a result appear quite rational but still be insane. . . . A man might know that he was doing wrong but be unable to stop himself (volitional insanity), or believe that he was not bound by the normal rules of society (emotional insanity).

103. SMITH, *supra* note 82, at 19.

104. *See* MODEL PENAL CODE § 4.01(1) (Proposed Official Draft 1962) ("A person is not responsible for criminal conduct if at the time of such conduct as a result of mental disease or defect he lacks substantial capacity either to appreciate the criminality [wrongfulness] of his conduct or to conform his conduct to the requirements of law." (alteration in original)).

was first introduced in *Parsons v. State*[105] and then famously adopted and developed by Judge David Bazelon in *Durham v. United States*.[106]

The *Durham* decision reversed the criminal conviction of Monte Durham, finding the trial court erred in holding that the defendant failed to raise sufficient proof to consider the insanity defense.[107] The court rejected the *M'Naghten* test and adopted a volitional approach that underscored the defendant's ability to choose to act in some way other than the way he did.[108] The proposed test read: "[A]n accused is not criminally responsible if his unlawful act was the product of a mental disease or mental defect."[109] There are two key elements in this test: the first is the soft definition of mental disease; the second is the hard determinism of the "products" test.

1. Soft Definition of Mental Disease

The *Durham* court made no attempt to define "mental disease" save that it addressed a condition that was capable of change in a way that a mental defect would not be.[110] The distinction did not prove helpful either to the courts or to psychiatrists called to testify because "[p]sychiatrists and some judges believed it established medicine's right to provide categories for classifying criminal deeds."[111] The lack of clarity is underscored by the historical event where psychiatrists redefined the term.[112] Immediately following the *Durham* decision, court-appointed psychiatric witnesses limited the definition of "mental disease" in insanity pleas to cases of psychosis because that was the standard for involuntary civil commitments at that time.[113] After a few years with this approach, mental health professionals changed the working definition of "mental disease" with no input from the courts. Professor Becker explains:

> [I]n 1957 . . . the staff of Saint Elizabeths Hospital decided to change its policy. Nonpsychotic diagnoses—particularly, the diagnosis of "sociopathic personality disturbance"—would now be explicitly recorded . . .[as] a "mental disease." . . .

105. 2 So. 854, 859 (Ala. 1887). I am grateful to Professor Bruce G. Berner of Valparaiso University School of Law for calling my attention to the original decision in which the volitional test emerged.

106. 214 F.2d 862, 874–75 (D.C. Cir. 1954).

107. *Id.* at 868–69.

108. *Id.* at 874–75.

109. *Id.*

110. *See id.*

111. SMITH, *supra* note 82, at 19.

112. *See generally* Loftus E. Becker, Jr., Durham *Revisited: Psychiatry and the Problem of Crime* (pt. 2), PSYCHIATRIC ANNALS, Sept. 1973, at 12 (1973).

113. *Id.* at 15.

> The change of policy at Saint Elizabeths had not been made as a result of any new psychological insights.[114]

This seemingly capricious shift undermined medical authority in the courts; suddenly the accepted standard changed, and none of the involved parties could explain why.[115]

2. Hard Determinism of the Products Test

The difficulty rooted in the nebulous definition of mental disease was compounded by its pairing with a fuzzy notion of causation in *Durham*'s "products test." After rejecting the *M'Naghten* test as inadequate, the court held that "an accused is not criminally responsible if his unlawful act was the product of mental disease or mental defect."[116] This test drew battle lines between lawyers committed to free will and psychiatrists who adopted a more determinist line. Psychiatrists preferred the *Durham* products test to *M'Naghten*'s line of authority. Professor Norrie observes:

> [T]hese tests . . . permitted a direct 'scientific' account of the accused's conduct to be delivered in cause and effect terms in the courtroom, unencumbered by old-fashioned and ultimately metaphysical tests of responsibility. *Durham* addressed the underlying causes of mentally disordered crime rather than dealing with what might only be certain symptoms of a disorder. It dealt concretely with the disordered subjectivity of the accused, and therefore was from the psychiatrists' viewpoint more just and understanding.[117]

In contrast to the shifting views of psychiatrists, lawyers and judges in the criminal justice system clung to the "old-fashioned . . . tests of responsibility."[118] Criminal lawyers labored daily to determine responsibility. Attorneys and judges wondered if the approach of psychiatry was at cross purposes with what they understood as the criminal law's primary function: assigning responsibility[119]:

> For lawyers, however, *Durham* represented a threat to the very notion of individual justice according to law. First, it took the decision out of the hands of both the law and the jury by making the question

---

114. *Id.* at 16. Professor Becker notes later that psychiatrists proposed this shift because their previous interpretation of the term seemed a far too restrictive definition of mental disease. *Id.* at 16. Becker sees this shift as a salutary maneuver because it permitted the psychiatrists to testify to other conditions which they had previously neglected to consider under the rubric of "mental disease," leading to more acquittals. *Id.* at 16–17. That said, the example described underscores the flimsiness of the definition and may add support to the legal profession's mistrust of forensic psychiatry. *See id.* at 17.

115. *See e.g.*, Alan A. Stone, *The Insanity Defense on Trial*, 33 HOSP. & COMMUNITY PSYCHIATRY 636, 637 (1982).

116. Durham v. United States, 214 F.2d 862, 874–75 (D.C. Cir. 1954).

117. NORRIE, *supra* note 100, at 183–84.

118. *Id.* at 184.

119. *Id.*

> of insanity a matter for psychiatry alone. The law was side-lined and the jury left with no real decision on the accused's responsibility. . . .
>
> The psychiatrists' scientific operating assumption of a universal *determinism* threatened to engulf the law's assumptions of free will and responsibility. . . . Scientific determinism was not a theory about insanity: it was a general theory about human conduct. Psychiatry threatened the liberal conception of the responsible subject.[120]

D. The Law in Dahmer

Wisconsin adopted the ALI test that combines both cognitive elements from the *M'Naghten* test and the volitional stress from *Durham*:

> [A defendant] is not responsible for criminal conduct if at the time of such conduct as a result of mental disease or defect the person lacked substantial capacity either to appreciate the wrongfulness of his or her conduct or conform his or her conduct to the requirements of law.[121]

In Wisconsin, a defendant must establish this affirmative defense to a reasonable certainty by the greater weight of the credible evidence.[122] Whether the defendant has met this burden of proof is a question of fact for the jury.[123] Of course, different jurisdictions may assign burdens of proof for affirmative defenses in different ways, but that consideration is beyond the scope of this paper's concerns.

V. CHALLENGES FACED BY THE DAHMER DEFENSE

A. The Challenge of Proving Insanity in General

As indicated above, criminal law assumes that punishment is appropriate only where a person can be blamed for his actions. For this reason, the law maintains that unless one's actions were the product of choice, we do not find guilt as a matter of law.[124] Because mental disease or defect can profoundly affect one's capacity to exercise rational choice, it seems that a defense ought be given in cases where a defendant's medical condition prevents its exercise.[125] But establishing that a medical or psychiatric condition exists can prove an elusive challenge. As the late

---

120. *Id.* (citation omitted).

121. *See* WIS. STAT. ANN. § 971.15(1) (West 2009).

122. *Id.* § 971.15(3).

123. *See* State v. Leach, 370 N.W.2d 240, 247 (Wis. 1985).

124. For example, courts have found that involuntary behavior due to psychomotor epilepsy may provide a defense to criminal liability. People v. Grant, 360 N.E.2d 809, 813 (Ill. App. Ct. 1977).

125. John J. McGrath, *The Insanity Defense: A Difficult Necessity*, 36 HOSP. & COMMUNITY PSYCHIATRY 54, 54–55 (1985):

> [T]he sane should be judged and sentenced, but the insane are proven to be ill and thus should be treated. It is crucial to distinguish between sanity and insanity during both judgment and disposition. . . . A return to Bedlam cannot be risked simply because our science finds the definition and demonstration of insanity difficult and elusive.

Professor Georges Canguilhem of the Sorbonne pointed out, mental disease, and therefore the insanity defense, is often difficult to determine.[126] His student Michel Foucault observed:

> [I]t is only by an artifice of language that the same meaning can be attributed to "illnesses of the body" and "illnesses of the mind." A unitary pathology using the same methods and concepts in the psychological and physiological domains is now purely mythical . . . .[127]

Psychiatric illnesses do not always have organic origins, and methods used in diagnosing and treating physical illnesses do not have clear parallels in psychiatry.[128] Conversely, it is also difficult to describe with precision what makes up a "normal" range of rational choice.[129] Thus, psychiatry lacks a clear standard to enunciate when a person ought to be held liable for their choices.[130] This apparent confusion stems from the uncertain organic basis for psychiatric maladies.[131] The causal factors linking body, mind, and behavior are still poorly understood.[132] The matter is further clouded by our positivist and empiricist bias, which largely holds that something does not exist unless it can be measured.[133] Mental and emotional problems are frequently not subject to empirical valida-

---

126. CANGUILHEM, *supra* note 42, at 117:

> Minkowski also thinks that the fact of insanity cannot be reduced to just the one fact of disease, determined by its reference to one image or precise idea of the average or normal [human] being. When we call another man insane, we do so intuitively "as men, not as specialists." The madman is "out of his mind" not so much in relation to other men as to life: he is not so much deviant as different.

127. MICHEL FOUCAULT, MENTAL ILLNESS AND PSYCHOLOGY 10 (Alan Sheridan trans., Univ. of Cal. Press 2d ed. 2008) (1954).

128. *Id.* at 10–11:

> [P]sychology has never been able to offer psychiatry what physiology gave to medicine: a tool of analysis that, in delimiting the disorder, makes it possible to envisage the functional relationship of this damage to the personality as a whole. The coherence of a psychological life seems, in effect, to be assured in some way other than the cohesion of an organism . . . . One cannot, then, make abstractions in the same way in psychology and in physiology, and the delimitation of a pathological disorder requires different methods in organic and in mental pathology.

129. *See* CANGUILHEM, *supra* note 42, at 119:

> In the final analysis it is the patients who most often decide—and from very different points of view—whether they are no longer normal or whether they have returned to normality. For a man whose future is almost always imagined starting from past experience, becoming normal again means taking up an interrupted activity . . . .

130. *See* Stone, *supra* note 115, at 640.

131. FOUCAULT, *supra* note 127, at 10–13.

132. CANGUILHEM, *supra* note 42, at 118 (noting how "somatic disease is capable of a superior empirical precision, of a better-defined standardization" than is mental disease).

133. *See, e.g.*, Martha C. Nussbaum, *Flawed Foundations: The Philosophical Critique of (a Particular Type of) Economics*, 64 U. CHI. L. REV. 1197, 1199 (1997) (noting how certain forms of economic analysis presume that all things can be measured and compared):

> A commitment to the commensurability of all an agent's ends runs very deep in the Law and Economics movement. Even when a plurality of distinct ends is initially recognized, the underlying view that agents are "maximizers of satisfactions," and that satisfaction is something that varies in degree rather than in kind, leads the theorist rapidly back to the idea that distinctions among options should be understood in terms of the quantity of utility they afford, rather than in terms of any basic qualitative differences.

tion, and psychiatrists observe that it is impossible to say if a patient is cured.[134]

The uncertain etiology of mental illness is often seen by jurists and the public as undermining the criminal justice system when it is invoked as the basis for excusing behavior.[135] Reflecting on the French experience, Foucault notes, "The essential issue therefore was to ascertain the reality and degree of the madness in question. The deeper the insanity, the more innocent the subject's will."[136] The public seems to believe that the insanity defense is successfully employed by large numbers of criminals who thereby avoid punishment.[137] Further, psychiatrists are mistrusted because they contradict each other on the stand.[138] Professor Alan Stone observed that the trial of John Hinckley, Jr., in which the defendant was acquitted because the prosecution failed to disprove mental illness beyond a reasonable doubt, "was a bleak experience for American psychiatry, and the verdict shook public confidence in the American criminal justice system."[139]

The above considerations underscore how the law and the public at large misperceive how psychiatrists can and do contribute to the legal enterprise. Initially, the prevailing clinical understanding of mental health issues is not easily translated into conclusions that can be of use in a courtroom.[140] This confusion in definition emerges in part because psychiatrists and attorneys have vastly different objectives when they inquire into psychiatric pathology. Clinical study of the mind is a therapeutic discipline. Psychiatrists and psychologists attempt to heal the suffering of those beset by mental and emotional distress. By contrast, lawyers focus on questions of blame and responsibility; they strive to divide those who have the ability to choose freely from those who cannot.[141] In cases raising the insanity defense, therefore, the questions to be grappled with at trial, though familiar ground for attorneys, are concerns far re-

---

134. Stone, *supra* note 115, at 640 ("Psychiatrists treat mental illness, often with great benefit to very sick patients, but that is not the same as curing them. . . . We can treat people and return them to the community. They will function better, but we cannot guarantee that they are cured . . . .").

135. CRONIN, *supra* note 88, at 93.

136. MICHEL FOUCAULT, HISTORY OF MADNESS 137 (Jean Khalfa ed., Routledge trans., 2006) (1961).

137. *See, e.g.*, Stone, *supra* note 115, at 637, where Professor Stone summarizes a lengthy statement by Hon. John Ashbrook of Ohio recorded in the 1981 Congressional Record.

138. *Id.*

139. ALAN STONE, LAW, PSYCHIATRY, AND MORALITY 77 (1984). John Hinckley, Jr. attempted to assassinate President Ronald Reagan in 1981. He was found not guilty by reason of insanity. *See* Stuart Taylor, Jr., *Hinckley Hails "Historical" Shooting to Win Love*, N.Y. TIMES, July 9, 1982, at A10, *available at* 1982 WLNR 315512.

140. CRONIN, *supra* note 88, at 90:

> The definition of insanity is a legal term, not a mental health term, and the defendant must meet the legal definition of being insane. The exact definition of insanity varies by jurisdiction. . . . Not everyone who suffers from a mental illness is judged by the courts to be insane. Indeed, many individuals who suffer from a psychosis and commit a crime do not meet the legal criteria for insanity.

141. *See* Stone, *supra* note 115, at 640.

moved from those of psychiatrists.[142] As a result, psychiatrists must be guarded in their testimony, and commentators worry that not all are.[143]

Not all challenges go against the defense; evidentiary rulings in insanity cases may cut both ways. Ordinarily, criminal cases are marked by strict limitations on admissible testimony. In the first instance, the evidence is restricted by the crime that has been charged.[144] The jury is directed to determine facts at a particular point in time.[145] Evidence of other actions or events is limited either by constitutional claims[146] or on the grounds of relevance.[147] Insanity trials broaden the frame of relevance; as the Wisconsin Supreme Court has noted, "no evidence should be excluded which reasonably tends to show the mental condition of the defendant at the time of the offense."[148] By pleading insanity, the defendant directs the jury to a question more amorphous than determining specific facts at a given point in time. Professor Wigmore is clear in his discussion of insanity pleas under common law: "Any and all conduct of the

142. *See* Insanity Def. Work Group, Am. Psychiatric Ass'n, *American Psychiatric Association Statement on the Insanity Defense*, 140 AM. J. PSYCHIATRY 681, 686 (1983):

> The American Psychiatric Association is not opposed to legislatures restricting psychiatric testimony . . . . We adopt this position because it is clear that psychiatrists are experts in medicine, not the law. As such, the psychiatrist's first obligation and expertise in the courtroom is to "do psychiatry," i.e., to present medical information and opinion about the defendant's mental state and motivation and to explain in detail the reason for his medical–psychiatric conclusions. When, however, "ultimate issue" questions are formulated by the law and put to the expert witness . . . then the expert witness is required to make a leap in logic. He no longer addresses himself to medical concepts but instead must infer or intuit what is in fact unspeakable, namely, the *probable relationship* between medical concepts and legal or moral constructs such as free will.

143. STONE, *supra* note 139, at 96:

> Psychiatry is held hostage by the psychiatrists who testify in courts whatever their standards and whatever the test of insanity may be. They undertake an enterprise which has hazards for us all. The reputation and credibility of our profession is in their hands. And if I am correct . . . they know not what they are doing.

144. FED. R. EVID. 402 ("All relevant evidence is admissible, except as otherwise provided by the Constitution of the United States, by Act of Congress, by these rules, or by other rules prescribed by the Supreme Court pursuant to statutory authority. Evidence which is not relevant is not admissible."). If the evidence proposed does not make a fact in consequence as determined by the substantive law more or less likely, it is not relevant. FED. R. EVID. 401. Therefore, courts streamline trials by excluding matters that do not bear directly on the case at hand, even limiting relevant evidence in some situations. *See* FED. R. EVID. 403.

145. RICOEUR, *supra* note 6, at 318:

> Past acts are . . . represented solely in terms of the nature of the charges selected prior to the actual trial. They are represented in the present within the horizon of the future social effect of the verdict that will decide the case. The relation to time is particularly noteworthy here: representation in the present consists in a staging . . . and a measured discourse of conscious legitimation . . . .

146. For example, under the Fifth Amendment, "No person . . . shall be compelled in any criminal case to be a witness against himself." U.S. CONST. amend. V. Relying in part on this provision, the U.S. Supreme Court held that defendants have a right to be informed of their rights to an attorney and to decline answering questions when subject to custodial interrogation. Miranda v. Arizona, 384 U.S. 436, 444 (1966).

147. FED. R. EVID. 403 ("Although relevant, evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury, or by considerations of undue delay, waste of time, or needless presentation of cumulative evidence.").

148. State v. Carlson, 93 N.W.2d 354, 361 (Wis. 1958).

person is admissible in evidence" as the jury attempts to determine if a defendant should be held accountable for his actions.[149] As Professor Goldstein observes:

> The almost unvarying policy of the courts has been to admit *any* evidence of aberrational behavior so long as it is probative of the defendant's mental condition . . . . Indeed, virtually never does one see any attempt to restrict the sort of lay evidence which is a staple of the insanity defense—that the defendant wept, or that he was given to violent rages, or that he threatened to throw his child out the window.[150]

By expanding the range of admissible evidence, the defendant pleading insanity has greater resources to mine than does the ordinary criminal defendant.[151] That said, when the defendant pleads guilty, the prosecution likewise can broaden the scope of the evidence it uses. Thus, the defense must be wary when entering a plea that throws open the doors to evidence ordinarily barred at trial.

B. Challenges Specific to the Dahmer Defense

The argument that anyone sexually attracted to corpses must have a mental disease seems self-evident, but that assertion faced specific difficulties in this case, both psychiatric and narrative. From the psychiatric perspective, no less an authority than Sigmund Freud doubted the claim that necrophiliacs suffer from a mental disease:

> Nevertheless, in some of these perversions the quality of the new sexual aim is of a kind to demand special examination. Certain of them are so far removed from the normal in their content that we cannot avoid pronouncing them 'pathological'. This is especially so where (as, for instance, in cases of . . . intercourse with dead bodies) the sexual instinct goes to astonishing lengths in successfully overriding the resistances of shame, disgust, horror or pain. But even in such cases we should not be too ready to assume that people who act in this way will necessarily turn out to be insane or subject to grave abnormalities of other kinds.[152]

Freud was quoted neither by the prosecution nor the defense in the Dahmer case, but his reluctance to find disease in such cases directly undermines the theory of the defense by observing that necrophiliacs can "override" resistances to their sexual acts and remain sane.[153]

Freud's objection points more generally to a logical flaw in the defense's strategy. The defense used circular logic to argue: (1) Jeffrey

149. 2 WIGMORE EVIDENCE § 228 (1979).
150. ABRAHAM S. GOLDSTEIN, THE INSANITY DEFENSE 54 (1967).
151. *Id.*
152. SIGMUND FREUD, *The Sexual Aberrations*, *in* THREE ESSAYS ON THE THEORY OF SEXUALITY 1, 27 (James Strachey ed. & trans., 1962).
153. *Id.*

Dahmer's sexual predilections were so disturbing that no sane person could share them, and (2) even if they could share them, they would not act on them; therefore, (3) because Dahmer had these urges and acted on them repeatedly, proves he could not control his actions. Therefore, he is insane, and his actions should be excused. When examined carefully, the position implies a wide range of disturbing conclusions. Change the facts a little. Assume rather that the defendant is sexually aroused only when he engages in acts of violent rape. Consistent with this deviation, he lures unsuspecting victims back to his apartment where he rapes them brutally. In his defense, he claims that he cannot control these urges; they are the only way he can achieve sexual satisfaction. Such a stimulus for sexual arousal is in many ways as distasteful as Dahmer's desires, but I doubt that most people share an intuition that the law should excuse the expression of violent rape fantasies. Merely because Dahmer had an unusual set of sexual triggers does not mean that he was less able to control himself than anyone else.[154] Indeed, Dahmer himself did not seem to subscribe to the "uncontrollable desire" argument:

> I have one person to blame—the person sitting across from you—no one else—no one put a gun to my head—I had choices to make and I made the wrong choices[.] I could have made different choices in the past[.] [I]t's obvious to me[.] If I had more foresight[,] if I had more motivation to find a career and worthwhile acts to fill my time—rather than drinking my problems away[.] I drank my emotions and problems away.[155]

The narrative difficulties faced by the defense are less direct and more complex. Initially, Dahmer laid out his actions in an extraordinarily detailed set of interviews with detectives and experts investigating the case. His statement to the Milwaukee Police Department alone fills over 145 typewritten pages.[156] Normally the defense controls the flow of information from the defendant; that was not the case here. Dahmer repeatedly asked to speak with officers, usually when his attorneys were present, but he sometimes insisted on speaking without counsel.[157] This

---

154. *See id.*

155. Frederick A. Fosdal, Interview Notes for Jeffrey Dahmer Examination 9 (Oct. 16, 1991) (unpublished report, on file with author).

156. *See* Statement of Jeffrey Dahmer, Case No. 2472 § 5 (unpublished police report, on file with author). Based on my experience as an assistant district attorney, this statement differs not only in quantity but also in kind. Before this case, I am hard pressed to recall a statement longer than perhaps ten typewritten pages. It is as though the psychic floodwaters came pouring out of the defendant once his dam of silence had been breached. In this way, it resembles the extensive narrative described with characteristic insight by Michel Foucault who saw the confession of the murderer Rivière and the dossier it engendered as expressing the battle among emerging professions trying to assert their power in evolving industrial society. This particular strand of thought is unfortunately beyond the scope of this paper. *See generally* MICHEL FOUCAULT, I, PIERRE RIVIERE, HAVING SLAUGHTERED MY MOTHER, MY SISTER, AND MY BROTHER . . . : A CASE OF PARRICIDE IN THE 19TH CENTURY (Michel Foucault ed., Frank Jellinek trans., Pantheon Books 1975) (1973).

157. *See, e.g.*, Dennis Murphy, Statement of Jeffrey Dahmer, Case No. 2472 § 5, at 47–48 (July 27, 1991) (unpublished police report, on file with author):

extensive confession hampered his attorneys' ability to craft a defense because they could not proffer any argument that conflicted with Dahmer's self-reported narrative.

A further difficulty in the case arose from its particularly gruesome facts, not only as reported by the defendant but also as photographed and collected by him.[158] On the one hand, the spoken and visual evidence could strengthen the argument that the defendant was mentally unhinged. On the other hand, the defense attorneys needed to weigh proffering evidence that could alienate the jury and risk a verdict based on disgust. Further, because of the physical evidence, the attorneys could not simply claim that Dahmer was delusional and made everything up. The physical evidence tied him ineluctably to facts reported. The defense therefore elected to clothe the evidence in a veneer of respectability by enveloping it in the testimony of clinicians. Concrete details of the murders and the disposal of the evidence were broadly "psychologized" so that the jury would focus on the predicament of a young man haunted by his unorthodox sexual urges, rather than looking at his bloodstained hands.

VI. MEETING THE CHALLENGES: NARRATIVE THEORY AND TRIAL COURTS

Trial attorneys are essentially storytellers, historians of brief moments in time who attempt to direct their audience to certain conclusions and not others.[159] Stories and historical accounts are usually enclosed in texts, a fixed set of symbols that mediate meaning from the author to the reader.[160] Although the trial court's decision in *Dahmer* was never appealed, and although no transcript was ever prepared, what occurred in

---

> I asked Mr. Dahmer if he had anything else to tell me, and why he did not request to have his attorney prior to talking to me, and he stated that he did not want his attorney there he just wanted to tell me about this other thing that he had forgotten and he didn't need his attorney present for that. I asked if he felt he needed the attorney for any other questioning, and he stated he felt he did not because he has been truthful with me the whole time and he does not feel he needs his attorney present when I'm there. I again informed him of his attorney's request, that he contact the attorney prior to contacting me, and he stated he understands and if he feels he has something important enough to tell me he will call me. I then informed him again that his attorney had requested his cooperation, and he stated he would consider it.

158. The court ordered sealed a set of Polaroid photos taken by the defendant after sentencing. I don't believe anyone has examined them since the trial.

159. *See, e.g.*, AMSTERDAM & BRUNER, *supra* note 5, at 110. *See generally* WHITE, CONSTITUTIONS AND RECONSTITUTIONS, *supra* note 5, at 3 (discussing the narrative nature of codified law); JAMES BOYD WHITE, *Telling Stories in the Law and in Ordinary Life: The* Oresteia *and "Noon Wine," in* HERACLES' BOW, *supra* note 5, at 168 (demonstrating the legal process as a story development and narrative method).

160. PAUL RICOEUR, HERMENEUTICS AND THE HUMAN SCIENCES: ESSAYS ON LANGUAGE, ACTION AND INTERPRETATION 174 (John B. Thomson ed. & trans., 1981):

> [I]n the asymmetrical relation between the text and the reader, one of the partners speaks for both. Bringing a text to language is always something other than hearing someone and listening to his speech. . . . For the text is an autonomous space of meaning which is no longer animated by the intention of its author; the autonomy of the text, deprived of this essential support, hands writing over to the sole interpretation of the reader.

the trial court can helpfully be understood as a text.[161] With attorneys in the authorial role and the jury cast as readers, witnesses spoke in front of the jurors but did not interact with them.[162] Jurors were not free to ask questions or to ask for clarification; rather, they were asked to apply the facts, as they found them, to the law.[163]

At a slightly different level of abstraction, the jury was asked to construct a master narrative, a world, a context for the evidence that emerged at trial and that somehow made sense of it.[164] The defense presented a series of events following in sequence that urged the jury to conclude that, because Jeffrey Dahmer's desires and actions were so bizarre, he must have been suffering from a mental disease. By contrast, the prosecution attempted to contextualize Dahmer's claimed madness by drawing a picture of Dahmer the man.[165] Ultimately, the members of

---

161. I use the term "text" here broadly. Legal scholars are used to referring to textual analysis at the appellate level. This makes perfect sense. Through printed texts and oral argument, appellate lawyers appeal to the rational and the propositional. By contrast, trial attorneys convey meaning through a much more diverse set of signifiers. In addition to being wordsmiths, trial attorneys need to employ the craft of a stage director or camera operator. Trial attorneys appreciate that juries will be moved by logic, but they also understand that reason is tutored by emotion. A bloodstained shirt, a wedding photo of the victim, or a weapon actually used in a murder adds little to the propositions that make up steps in a chain of reasons. However, these mute objects may speak eloquently of the ebb and flow of a homicide victim's life, the terror of witnesses, and the burden of sorrow carried by family members.

162. RICOEUR, *supra* note 160, at 146–47:

> It does not suffice to say that reading is a dialogue with the author through his work, for the relation of the reader to the book is of a completely different nature. Dialogue is an exchange of questions and answers; there is no exchange of this sort between the writer and the reader. The writer does not respond to the reader. . . . [Reading] thereby replaces the relation of dialogue, which directly connects the voice of one to the hearing of the other.

163. *See* PAUL RICOEUR, THE JUST 121 (David Pellauer trans., Univ. of Chi. Press 2000) (1995):

> The application of a rule is in fact a very complex operation where the interpretation of the facts and the interpretation of the norm mutually condition each other, before ending in the qualification by which it is said that some allegedly criminal behavior falls under such and such a norm which is said to have been violated. If we begin with the interpretation of the facts, we cannot overemphasize the multitude of ways a set of interconnected facts can be considered and, let us say, recounted. . . . We never finish untangling the lines of the personal story of an accused with certainty, and even reading it in such a way is already oriented by the presumption that such an interconnectedness places the case under some rule. To say that *a* is a case of *B* is already to decide that the juridical syllogism holds for it.

164. RICOEUR, *supra* note 160, at 178:

> What we make our own, what we appropriate for ourselves, is not an alien experience or a distant intention, but the horizon of a world towards which a work directs itself. The appropriation of the reference is no longer modelled on the fusion of consciousnesses, on empathy or sympathy. The emergence of the sense and reference of a text in language is the coming to language of a world and not the recognition of another person.

165. *Id.* at 278:

> [I]t must be said that any narrative combines, in varying proportions, two dimensions: a chronological dimension and a non-chronological dimension. The first may be called the 'episodic dimension' of the narrative. Within the art of following a story, this dimension is expressed in the expectation of contingencies which affect the story's development; hence it gives rise to questions such as: and so? and then? what happened next? . . . . But the activity of narrating does not consist simply in adding episodes to one another; it also constructs meaningful totalities out of scattered events. This aspect of the art of narrating

the jury had to reconcile these two approaches in terms of the law they were given. They had to decide which narrative, if either, corresponded with what they understood as the truth.[166]

As Paul Ricouer observes in his work on nonfiction narratives, historical truth as expressed in texts is always a constructed entity; even if it does not correspond with the historical events it purports to describe, it may well supplant the event itself in the community's imagination.[167] Historians understand this distinction as a matter of course. Professor Tzvetan Todorov observes:

> The work of the historian, like every work on the past, never consists solely in establishing the facts but also in choosing certain among them as being more salient and more significant than others, then placing them in relation to one another . . . .[168]

Just as the historian chooses among salient facts, attorneys select what evidence to put before the jury.[169] Thus, it matters greatly what evidence is brought before the jury, for that testimony is the only basis whereby the jury can construct its sense of "what really occurred."[170] Of

---

is reflected, on the side of following a story, in the attempt to 'grasp together' successive events. The art of narrating, as well as the corresponding art of following a story, therefore require that we are able *to extract a configuration from a succession.*

166. RICOEUR, *supra* note 6, at 178–79:

> A vigilant epistemology will guard here against the illusion of believing that what we call a fact coincides with what really happened, or with the living memory of eyewitnesses, as if the facts lay sleeping in the documents until the historians extracted them. This illusion . . . for a long time underlay the conviction that the historical fact does not differ fundamentally from the empirical fact in the experimental natural sciences. . . . [W]e need to resist this initial confusion between a historical fact and a really remembered event. The fact is not the event, itself given to the conscious life of a witness, but the contents of a statement meant to represent it. . . . So understood, the fact can be said to be constructed through the procedure that disengages it from a series of documents concerning which we may say in return that they establish it. This reciprocity between construction (through a complex documentary procedure) and the establishing of a fact (on the basis of the document) expresses the specific epistemological status of the historical fact. It is this propositional character of the historical fact (in the sense of "fact that . . .") that governs the mode of truth or falsity attached to the fact. The terms "true" and "false" can legitimately be taken at this level in the Popperian sense of "refutable" and "verifiable."

(third omission in original).

167. *Id.*

168. *Id.* at 86 (internal quotation marks omitted) (quoting TZVETAN TODOROV, LES ABUS DE LA MÉMOIRE 50 (1995).

169. *Id.* at 318:

> Past acts are . . . represented solely in terms of the nature of the charges selected prior to the actual trial. They are represented in the present within the horizon of the future social effect of the verdict that will decide the case. The relation to time is particularly noteworthy here: representation in the present consists in a staging, a theatricalization . . . . This living presence of the scenes replayed solely on the plane of discourse comes under the heading of visibility, which was shown . . . to be related to the expressibility . . . on the plane of the literary representation of the past.

170. CHATMAN, *supra* note 8, at 45–46:

> [T]he interesting thing is that our minds inveterately seek structure, and they will provide it if necessary. Unless otherwise instructed, readers will tend to assume that even "The king died and the queen died" presents a causal link, that the king's death has something to do with the queen's. We do so in the same spirit in which we seek coherence in the

course, the jury's reconstruction, based on its limited information, may differ greatly from the historical event itself.[171] Despite this variance from the historical event, in our system of justice, the jury's verdict is the version of history that matters. Thus, the defense crafted an object of discourse, of words, to supplant the flesh and blood reality of Jeffrey Dahmer, leading the jury to focus on only parts of his story to persuade them to adopt a version of events at odds with the broader historical narrative. They did this by creating two claims of consistency that sublimated Dahmer's normality and attempted to hide his guilt.

The substance or subject matter of a narrative, or trial, cannot be separated from its medium, how the story is told, and who tells the jury what it hears.[172] In Dahmer's case in chief, the jury saw his life primarily through the lens of detectives, clinical psychologists, and psychiatrists. This narrative distance conferred a respectability upon the content which would have been absent had Dahmer been testifying on his own behalf.[173] If Dahmer were to recount the same events from the witness stand, the jurors would naturally regard his testimony as self-serving and discount its possible truthfulness.[174] By contrast, when that same testi-

---

visual field, that is, we are inherently disposed to turn raw sensation into perception. So one may argue that pure "chronicle" is difficult to achieve. "The king died and then the queen died" and "The king died and then the queen died of grief" differ narratively only in degrees of explicitness at the surface level; at the deeper structural level the causal element is present in both. The reader "understands" or supplies it; he infers that the king's death is the cause of the queen's. "Because" is inferred through ordinary presumptions about the world, including the purposive character of speech.

171. RICOEUR, *supra* note 6, at 179:

> What is one talking about when one says that something happened? . . . [I]t is to preserve this status of the reference of historical discourse that I distinguish the fact as "something said," the "what" of historical discourse, from the event as "what one talks about," the "subject of . . ." that makes up historical discourse. In this regard, that assertion of a historical fact indicates the distance between the said (the thing said) and the intended reference, which according to one of Benveniste's expressions turns discourse back toward the world. The world, in history, is past human life as it happened. . . . [What is said is known as] "standing for" . . . . To get there, we need to leave underdetermined the question of the actual relation between fact and event, and tolerate a certain indiscrimination in the employment by the best historians of these terms as standing for each other.

172. *Id.* at 163–64:

> The specificity of testimony consists in the fact that the assertion of reality is inseparable from its being paired with the self-designation of the testifying subject. The typical formation of testimony proceeds from this pairing: I was there. What is attested to is indivisibly the reality of the past thing and the presence of the narrator at the place of its occurrence. And it is the witness who first declares himself to be a witness. . . . These . . . assertions link point-like testimony to the whole history of a life.

(citation omitted).

173. *See id.*

174. Indeed, the jury instructions in Wisconsin would explicitly invite this sort of reasoning on the jury's part. *See, e.g.*, WIS. JI–CRIMINAL § 300, at 1 (2003):

> It is the duty of the jury to scrutinize and to weigh the testimony of witnesses and to determine the effect of the evidence as a whole. You are the sole judges of the credibility, that is, the believability, of the witnesses and the weight to be given to their testimony.
>
> In determining the credibility of each witness and the weight you give to the testimony of each witness, consider these factors:

mony comes through the mouth of a psychiatrist or clinical psychologist, the jury attaches unwarranted credibility to the claims made.[175] This is not to say that the jury is necessarily persuaded by mental health professionals or police officers, but still, the defendant's words seem less self-interested when audible in their voices rather than his own.

The effects of the gap between Dahmer and those who recounted his statements are seen in cross examination. The prosecutors typically addressed not the internal consistency of Dahmer's statements, but the prudential judgments made by the clinician in light of the evidence.[176] By encasing Dahmer's statements within sworn testimony of medical professionals, the defense largely insulated the substance of Dahmer's statements from challenge. Although the jury might have questioned the clinical witnesses' conclusions, it did not consider directly the truthfulness of Dahmer's self-report.[177]

The failure to interrogate Dahmer's statements for truthfulness is not insubstantial. The defense repeatedly asserted that it was being trans-

---

- whether the witness has an interest or lack of interest in the result of this trial;
- the witness' conduct, appearance, and demeanor on the witness stand;
- the clearness or lack of clearness of the witness' recollections;

. . . .

- bias or prejudice, if any has been shown;
- possible motives for falsifying testimony; and
- all other facts and circumstances during the trial which tend either to support or to discredit the testimony.

175. RICOEUR, *supra* note 6, at 164:

> It is before someone that the witness testifies to the reality of some scene of which he was part of the audience, perhaps as actor or victim, yet, in the moment of testifying, he is in the position of a third-person observer with regard to all the protagonists of the action. This dialogical structure immediately makes clear the dimension of trust involved: the witness asks to be believed. He does not limit himself to saying "I was there," he adds "believe me." Certification of the testimony then is not complete except through the echo response of the one who receives the testimony and accepts it. Then the testimony is not just certified, it is accredited. . . . In this case, the accreditation comes down to authenticating the witness on personal terms. The result is what we call his trustworthiness, whose evaluation can be assimilated to comparative orders of magnitude.

(citation omitted).

176. *Id.* at 164–65:

> The possibility of suspicion in turn opens a space of controversy within which several testimonies and several witnesses find themselves confronted with one another. . . . The witness anticipates these circumstances in a way by adding a third clause to his declaration: "I was there," he says, "believe me," to which he adds, "If you don't believe me, ask someone else," said almost like a challenge. The witness is thus the one who accepts being questioned and expected to answer what may turn out to be a criticism of what he says.

177. The testimony of the court's own psychiatric expert, Dr. George Palermo, addresses this very issue. *See* David Doege, *Anger in His Homosexuality Led Dahmer to Kill, Psychiatrist Says*, MILWAUKEE SENTINEL, Feb. 7, 1992, at 1:

> Dahmer has lied for years and still lies today.
>
> "He lied to the judge in 1989 (when Dahmer was sentenced for sexual assault)," Palermo said. "He lied to his lawyer."
>
> "He lied to many doctors to get the (sleeping) pills. It is my feeling he has embellished a great deal in the things he has said he did."

parent, when it was actually casting significant parts of the story into the shadows.[178] For example, woven into the testimony of the psychiatrists was an unstated assumption that Dahmer's statements were at all times internally consistent; that he should be believed because he unburdened himself completely and uniformly. The evidence does not support that assertion.[179]

Indeed, there are wide factual disparities, both large and small, in Dahmer's recounting of the murder of Stephen Hicks, the first murder he admitted committing. In his initial statement, Dahmer relates that he and Hicks had sex at his home and later fought, and Hicks died, almost by accident, when Dahmer struck him in the head with a barbell.[180] The story develops over time. When he next describes the Hicks homicide, he states that Hicks was not a homosexual, and they did not have sex and does not mention a fight.[181] In a later statement, he underscores that he and Hicks did not engage in any homosexual activities before or after Hicks's death.[182] When talking with one of the examining psychiatrists, Dahmer related that he "[h]ad the idea to hit [Hicks] over the head for

178. RICOEUR, *supra* note 6, at 85 ("It is, more precisely, the selective function of the narrative that opens to manipulation the opportunity and the means of a clever strategy, consisting from the outset in a strategy of forgetting as much as in a strategy of remembering.").

179. *Id.* at 316–17:

> To be sure, applying the criteria of concordance and relying upon independent verification of the confession provide perfect illustrations of the theses offered . . . on the "evidentiary paradigm": the same complementarity between the oral nature of testimony and the material nature of the evidence authenticated by expert testimony; the same relevance of "small errors," the probable sign of inauthenticity; the same primacy accorded to questioning, to playing with possibilities in imagination; the same perspicacity in uncovering contradictions, incoherencies, unlikelihoods; the same attention to silences, to voluntary or involuntary omissions; the same familiarity, finally, with the resources for falsifying language in terms of error, lying, self-delusion, deception.

180. Patrick Kennedy, Statement of Jeffrey Dahmer, Case No. 2472 § 5, at 9–10 (July 23, 1991) (unpublished police report, on file with author):

> Subject [Jeffrey Dahmer] states that when he was 18 years of age and living in Richfield, Ohio, he picked up a hitchhiker whom he described as a white male about 19 years of age. He states he took him home and had homosexual sex with him and states they were drinking beer and became intoxicated. He states they got into a physical fight because the 19-year old individual tried to leave and that during the fight, he states he struck the hitchhiker with a barbel [sic]. He states that the blow of the barbel [sic] caused the death of the hitchhiker, and at this time he took the body out into a wooded area by his house and left it there to decompose for about two weeks.

181. Dennis Murphy, Statement of Jeffrey Dahmer, Case No. 2472 § 5, at 23 (July 24, 1991) (unpublished police report, on file with author):

> He states his first homicide which occurred, he believes around October of 1978, was of a white male hitchhiker, whom he describes as 18-yoa., 5'10" tall, skinny build, maybe 150 lbs., having straight brown collar length hair, not wearing glasses, clean shaven, and he believes he was not a homosexual. He states he didn't have sex with this individual, he just invited him in for a drink, and when the individual wanted to leave that's when he hit him with a "barbell" and subsequently disposed of the body behind his residence. He states he did burn his clothes and identification.

182. R.W. Munsey, Statement of Jeffrey L. Dahmer, Case No. 2472 § 12, at 488 (Aug. 8, 1991) (unpublished police report, on file with author) ("Jeffrey L. Dahmer was interviewed regarding whether he engaged in any homosexual activities with Steven M. Hicks before or after his death. Jeffrey L. Dahmer stated there were 'no' homosexual activities.").

about a half hour."[183] He then strangled him with a barbell, explaining, "I didn't want to get caught so I went all the way and finished it (strangled him)."[184] The day after the murder, Dahmer continued, he masturbated in front of the body and touched the body in the chest and penis area, but did not have oral sex.[185] The following day he bought a hunting knife, masturbated again, cut him open to view his insides, and then dismembered the body "to make him light enough to carry."[186] In a later interview with a different psychiatrist, Dahmer claims he opened Hicks's belly and masturbated over that and later cut his head off, cleaned it off under the sink, and masturbated in front of that.[187]

Had Dahmer testified at trial, each of these inconsistencies would have been laid out in detail before the jury to cast doubt on his veracity. Because the evidence came in through clinicians and detectives though, this line of cross examination was never developed. Therefore, the contradictions within Dahmer's statements were largely ignored, except for Dr. Palermo's assertion that Dahmer "embellished" much of what he claimed to have done,[188] and the jury was given the impression that Dahmer was basically truthful.

It is noteworthy that Jeffrey Dahmer's statements describing the Hicks murder grew more bizarre and disturbing as the trial date approached. By this time, Dahmer had likely internalized a desire to be found psychologically ill rather than wicked. Because Dahmer's statements furnish the sole source of information given to the jury, his motives to testify falsely at trial needed to be fully explored. In his varying renditions of the Hicks murder, Dahmer may have reflexively shaded the truth to appear more psychologically ill than he was.

The ostensible internal consistency of Dahmer's statements projected by the defense echoes a coordinated assertion that Dahmer's actions in the fifteen charged murders were similar. Rather than breaking down each murder to explore their unique circumstances, the defense implied that Dahmer's killings were locked in a repeating loop. Although the defense addressed the specifics of the uncharged Hicks and Tuomi murders, the fifteen charged offenses were grouped as an indistinguisha-

183. Frederick A. Fosdal, Interview Notes for Jeffrey Dahmer Examination 13 (Oct. 23, 1991) (unpublished report, on file with author).

184. *Id.*

185. *Id.*

186. *Id.* at 14.

187. Park Elliott Dietz, Interview Notes for Jeffrey Dahmer Examination, at A81–A82 (Jan. 5, 1992) (unpublished report, on file with author).

188. *See, e.g.*, Jim Stingl, *Dahmer Needs Help but Is Sane, Court Told*, MILWAUKEE J., Feb. 6, 1992, at 1 (internal quotation marks omitted):

> [Dr.] Palermo called Dahmer a manipulator and said he doubted Dahmer's claim that he planned to build a temple out of the bones of his victims.
>
> . . . .
>
> He also doubted whether Dahmer actually ate the flesh of any of his victims. Dahmer "embellished" the facts and made them more ugly than they already were, he said.

ble whole by Dahmer's attorneys and expert witnesses.[189] For example, Dr. Fred Berlin of the Johns Hopkins Medical School focused on the process of differential diagnosis in reaching his conclusion that the defendant suffered from a mental disease, rather than highlighting facts of the crimes related by the defendant.[190] After describing Dahmer's disappointment and hopelessness following the death of Steven Tuomi, Berlin concluded that "Mr. Dahmer was out of control. . . . It wasn't going to be he who was going to stop it. It was going to have to be an outside force."[191] Dr. Berlin barely mentioned the name of any of Dahmer's other victims on direct examination.[192]

Similarly, Dr. Judith Becker developed a narrative regarding Dahmer's developmental psychology. She brought out childhood memories and tried to tie them to his later actions.[193] She testified that Dahmer was eventually so "consumed" by the mental disease of necrophilia that his obsession led him to kill uncontrollably.[194] Becker's testimony revealed some differences from Dahmer's statement to the police; he stated to her that he had sex with every victim's body after death.[195] Becker also discussed specifics of the murders themselves, including a new revelation

---

189. Videotape: State of Wisconsin v. Jeffrey Dahmer Tape One, at 46:10 (Jan. 30, 1992) (on file with the Marquette University Law Library). In his opening statement, Dahmer's attorney pointed out that, in the opinion of his three expert witnesses, after Dahmer's killing of Steven Tuomi in the Ambassador Hotel in Milwaukee, "It was all over. Mr. Dahmer would continue to do this until it was stopped." *Id.*

190. Videotape: State of Wisconsin v. Jeffrey Dahmer Tape Three, at 2:59:45 (Feb. 3, 1992) (on file with the Marquette University Law Library). Dr. Berlin described psychiatric illness as embodying a value judgment. *Id.* He stated that there is a diversity of bodily conditions and posed the following argument: In physical medicine, consider two different sorts of conditions or processes. One is called cancer, the other is called respiration or breathing. We don't like cancer. It causes suffering; therefore, we call it a disease. Similarly, there are different sexual attractions. Some people are heterosexually attracted; others are attracted to persons of the same gender. At one point, psychiatry thought that homosexuality was a disease. It no longer does. Dahmer's case is different. He has intense, recurrent sexual fantasies of dead bodies. *Id.*; *see also Testimony from the Dahmer Trial*, MILWAUKEE SENTINEL, Feb. 4, 1992, at 6 ("If this isn't mental illness, from my point of view, I don't know what is.").

191. David Doege, *Doctor: Dahmer 'Out of Control' Testifies on Killer's Fantasies*, MILWAUKEE SENTINEL, Feb. 4, 1992, at 1 (internal quotation marks omitted).

192. *See id.* On cross examination, the prosecutor elicited that Dr. Berlin spent only four hours and forty-five minutes total with Dahmer in reaching his diagnosis, less than half the time spent by any other expert witness. *See id.* This information may have diminished Dr. Berlin's effectiveness in the jury's eyes.

193. *Testimony from the Dahmer Trial*, MILWAUKEE SENTINEL, Feb. 5, 1992, at 7:

> At age 8 . . . he recalled his father . . . fished in the pond . . . .
>
> . . . . When he talked of cutting the fish open and seeing the inside of the fish, he became somewhat more animated, somewhat more alive in a sense.
>
> Knowing what he had done to his victims and cutting them open and knowing that he had appeared to be fascinated by the viscera, by the insides of his victims, I wonder if that early incident of the cutting open and the fascination with colors was not somehow related to what happened later on.

(first and second omission in original).

194. Doege, *supra* note 41, at 1.

195. *Id.*

that he had intercourse with the internal organs of his victims.[196] She also discussed other gruesome aspects of the case such as the cannibalism of Ernest Miller and Errol Lindsey.[197] That said, Dahmer's careful planning and efforts to avoid capture were absent from her testimony.[198] Dr. Becker's testimony was an interpretation of Dahmer's actions from a psychological model rather than a description of his actions; she wove a narrative of what might have been going on in Dahmer's mind rather than describing the whole of what he did.[199] Her failure to wrestle with Dahmer's actions led her to minimize obvious aspects such as the need to dispose of bodies before they started rotting. Dr. Becker also spent a great deal of time explaining Dahmer's current suicidal ideation and his plans for a bone shrine he never built.[200]

Dr. Becker's fascination with the psychological narrative is demonstrated by her willingness to accept Dahmer's description of his proposed "temple of bones," which did not match up with the physical evidence in the case. Dahmer sketched for her a diagram of what the "temple" would look like.[201] It would be built around a black lacquered table that would have two skeletons on either side of a desk resting at hip height, and there would be skulls he collected on the table looking back at him.[202] He claimed he had already purchased the table.[203] The difficulty emerges when one considers the physical evidence. The only table Dahmer owned was a black coffee table about 15 to 18 inches tall and approximately four feet long; it appeared in some of the photos of his victims. None of the skeletons he collected could rest on it at hip height because it was not tall enough. Further, the skulls could not be stacked in the ways he claimed because the surface was too small to hold them. Dr. Becker simply failed to check Dahmer's claims against evidence on police inventory.[204] Similarly, a third clinician, Dr. Wahlstrom, focused on Dah-

196. Jim Stingl, *Several Disorders Played Role: Expert Dahmer Was Psychotic, Witness Testifies*, MILWAUKEE J., Feb. 5, 1992, at 1. There is no evidence that Dahmer mentioned this act of masturbating with internal organs to any other witness. Indeed, it is quite clear he did not say this to either Dr. Fosdal or Dr. Dietz. Both of them spoke with the defendant after Dr. Becker's interview. I find it difficult to account for Dahmer's failure to mention this particular paraphilia to other witnesses when he had admitted so much already.

197. *Testimony from the Dahmer Trial*, *supra* note 193, at 7.

198. *Id.*

199. Dr. Becker acted as an excellent psychologist in trying to describe Dahmer's mental processes. However, her analysis is not how the law looks at intent. As James Fitzjames Stephen points out, "The only possible way of discovering a man's intention is by looking at what he actually did, and by considering what must have appeared to him at the time the natural consequence of his conduct." STEPHEN, *supra* note 59, at 111. By foregrounding the psychological history, Dr. Becker de-emphasized the physical facts of the case.

200. *See* Doege, *supra* note 41, at 1.

201. *See id.*

202. *See* Stingl, *supra* note 196, at 1.

203. *See id.* ("Dahmer said the temple, to be built on a black table in his West Side apartment, would be a 'power center.'").

204. Indeed, the court-appointed expert, Dr. George Palermo, "doubted Dahmer's claim that he planned to build a temple out of the bones of his victims." Stingl, *supra* note 188, at 1. It may have

mer's delusions regarding his possible temple of bones.[205] Wahlstrom's testimony in particular dwelt on Dahmer's performance on a battery of psychological tests.[206] However, his testimony, like that of Drs. Berlin and Becker, did not delve much into the particulars of the crimes themselves.[207] In the final analysis, his opinions seemed based largely on broad abstractions, and he therefore failed to provide the jury with concrete evidence to refute the state's case.

Dahmer's failure to testify impeded another possible line of questioning: had Dahmer spoken under oath, the prosecution could have developed how he enticed men to come home with him. In the case-in-chief, Dahmer rendered these encounters as largely financial.[208] What if this report was another of Dahmer's manipulations? It may be that one of the most important ways Dahmer showed control was in his ability to appear attractive, friendly, and a safe person with whom to go home. Because he did not discuss these matters with the detectives or expert witnesses, the jury could not reflect on this aspect of his personality. From one perspective, his silence in the courtroom may have been that of a puppeteer: he provided the words to the witnesses that he could not say himself. By insulating himself from questions that would shatter his claim of illness, he may have been manipulating the defense witnesses just as he manipulated his victims.

At the end of the defense case-in-chief, the jury had a description of a disturbed figure who did horrifying things. Surely that is one aspect of Jeffrey Dahmer. However, he was much more. The image crafted by the defense failed to account for how Dahmer could hold a job or persuade his victims to return to his apartment with him. The defense focused the jury on the bizarre and delusional, asking it to return a verdict based on these disjointed episodes in the life of the silent man surrounded by his attorneys at counsel table. It was not clear at the end of the defense's case how someone as disturbed as the Jeffrey Dahmer they had presented managed to function undetected for years.

VII. MEETING THE CHALLENGES: REFRAMING A MADMAN AS A CONSUMMATE PLANNER

To counter the discursive image of a madman suggested by the defense, the prosecution filled in the picture of the defendant, showing that

---

stemmed from the failure of Dahmer's story to match up with the physical evidence, but the record does not address that point.

205. Stingl, *supra* note 196, at 1 ("Carl M. Wahlstrom Jr., who became a psychiatrist two years ago and works in Chicago, said he believed Dahmer was psychotic because of his plans to erect a temple where he would display the skulls and bones of his victims.").

206. *See* Carl M. Wahlstrom, Jr., Report of Psychiatric Evaluation of Jeffrey Dahmer 9–10 (Jan. 9, 1992) (unpublished court file, on file with author).

207. In an eleven-page report, brief descriptions of the crimes themselves appear in only four paragraphs. *See id.* at 6–7.

208. *See, e.g.*, Murphy & Kennedy, *supra* note 29, at 88.

he was at all times in control of his actions. As a segue into the prosecution's case-in-chief, the court's witness, Dr. George Palermo, maintained that Jeffrey Dahmer was responsible for his actions under the law.[209] Palermo described Dahmer as making conscious choices "at the moment of the killings, in the preparation of the killing and afterwards."[210] Palermo anticipated the prosecution's strategy by stating "Jeffrey Dahmer is a human being . . . . To take [that] away from him by just saying he is a necrophile is wrong . . . . He is much more."[211] He continued, "Jeffrey Dahmer knew exactly what he was doing . . . he had taken precautions, very, very good ones. He knew the consequences of his action but he did not want to stop."[212]

The prosecution took as its theme Dahmer's desire to control and his ability to choose, and it reframed earlier testimony in the case as indicating careful planning on Dahmer's part.[213] Initially, the prosecution needed to normalize Jeffrey Dahmer, to provide a context, by showing that his life fit together coherently; he struck others as friendly, unremarkable, and sane.[214] Beginning with the cross-examination of Detective Dennis Murphy, the prosecution asked every witness who had exten-

---

209. Stingl, *supra* note 188, at 1:

> Palermo said that like most people, he expected to come face-to-face with a "crazy" person the first time he met with Dahmer, because of the number of people Dahmer had killed.
>
> "I was shocked when I met him," Palermo said. "I knew after four hours that he was not psychotic."
>
> Palermo said Dahmer's speech was clear and his answers were coherent. He found Dahmer amiable and intelligent . . . . "He's a likable fellow."

210. *Id.* (internal quotation marks omitted).

211. Doege, *supra* note 177, at 1 (internal quotation marks omitted).

212. *Testimony from the Dahmer Trial*, MILWAUKEE SENTINEL, Feb. 7, 1992, at 7 (alteration in original).

213. *See, e.g.*, Jim Stingl, *Urge to Kill Ruled: Expert*, MILWAUKEE J., Feb. 3, 1992, at 1. As Detective Murphy noted:

> Dahmer felt a sense of power knowing his family, neighbors and even police officers couldn't detect his secret world of killing. "He took pleasure in the fact of knowing that he had a private world of his own that no one else knew about," Murphy said. "He felt he had this ability to make people see a phase of him that only he wished them to see, and this encouraged him to continue on with his crimes, feeling that he would never be caught," Murphy said.

Id.

214. Under *Duthey v. State*, the jury is permitted to rely on the opinions of laypersons who are familiar with the defendant about his sanity or lack thereof. 111 N.W. 222, 226 (Wis. 1907). The substance of this common law decision is encased in the Federal Rules of Evidence on opinion testimony by lay witnesses:

> If the witness is not testifying as an expert, the witness' testimony in the form of opinions or inferences is limited to those opinions or inferences which are (a) rationally based on the perception of the witness, (b) helpful to a clear understanding of the witness' testimony or the determination of a fact in issue, and (c) not based on scientific, technical, or other specialized knowledge within the scope of Rule 702.

FED. R. EVID. 701. The witness has to have first-hand knowledge, and this knowledge must be of the sort that will help the jury resolve a disputed fact. This was the foundation that the prosecution used in introducing testimony from Dahmer's co-workers, apartment manager, police officers, and potential victims who testified as to Dahmer's apparent sanity.

sive contact with Jeffrey Dahmer the same set of questions. After establishing that the witness had spent an appreciable amount of time with Dahmer, the examiner would ask if Dahmer displayed hallucinations, delusions, unconnected thoughts, incoherent responses, or appeared not to be tracking the conversation.[215] In this manner, the prosecution built up a weight of evidence from ordinary people who encountered the defendant in various times and places, and who saw him as perfectly normal and unremarkable.

For example, his former boss at the Ambrosia Chocolate Company testified that he had no problems with Dahmer, whom he described as "polite."[216] "He was quiet. He had no problems reacting with others."[217] Further, his boss thought Dahmer did "a satisfactory job."[218] One of the facts that came out at trial was that Dahmer was able to mix almost five hundred distinct chocolate recipes during his time there, indicating his ability to perform and be paid for complex tasks. Dahmer's apartment manager thought Dahmer was "a very nice guy," and he was willing to ask Dahmer to become his business partner.[219] His building manager further described Dahmer's apartment as "probably the neatest apartment I've seen."[220] This line of questioning helped the jury see Dahmer as he was seen by coworkers and other ordinary people. He did not strike acquaintances as out of touch with reality.

A second sort of lay witness helped the jury focus on Dahmer's mental state at or near the time of his attempted or completed murders. The prosecution called citizen witnesses and police officers, whom had observed Dahmer near these times, to testify to his apparent rationality and control. A friend of Dahmer's fifth victim, Anthony Sears, explained he dropped Dahmer and Sears off near Dahmer's grandmother's home on the last night he saw his friend.[221] "I felt that [Dahmer] was a very nice person. He seemed very kind."[222] Another witness, Ronald Flowers, had difficulties starting his car, so Dahmer offered to take him to his grand-

---

215. *See, e.g.*, Videotape: State of Wisconsin v. Jeffrey Dahmer Tape Three, at 48:20 (Feb. 3, 1992) (on file with the Marquette University Law Library). Detective Dennis Murphy denied on cross examination that Dahmer displayed any evidence of mental illness during the approximately sixty hours of interviews he had with him. *Id.* He further indicated that Dahmer was able to describe in detail the extent of his planning, the lengths he went to in eliminating the evidence, and that he felt he had the ability to make people see only what he wanted. *Id.*

216. David Doege, *Dahmer's Work OK, Bosses Say*, MILWAUKEE SENTINEL, Feb. 11, 1992, at 1.

217. *Id.* (internal quotation marks omitted).

218. *Id.* (internal quotation marks omitted).

219. *Testimony from the Dahmer Trial*, MILWAUKEE SENTINEL, Feb. 12, 1992, at 8.

220. *Id.*

221. Jim Stingl, *Dahmer Is Sane, Psychiatrist Says*, MILWAUKEE J., Feb. 8, 1992, at 1.

222. *Id.* (internal quotation marks omitted). This observation goes to Dahmer's ability to be manipulative. He could be charming and attractive to other men, a fact largely unexplored during trial.

mother's home to pick up jumper cables.[223] After the cab dropped them off near Dahmer's address, Flowers testified that he was suspicious and warned Dahmer he only wanted to get his car started.[224] Flowers came inside the home reluctantly, and Dahmer said that he was tired and needed some coffee. Flowers agreed and passed out soon after drinking the coffee mixed with Halcion that Dahmer had prepared for him.[225] The next thing Flowers remembered was awaking in a hospital room.[226] Flowers stated that he encountered Dahmer a year later in a bar; Dahmer said to him, "I really don't remember who you are . . . maybe we can go have a cup of coffee."[227]

On May 27, 1991, Jeffrey Dahmer lured a young Laotian man, Konerak Sinthasomphone, back to his apartment. He claimed to have drugged him, drilled a hole in his skull, and injected him with a dose of muriatic acid and water.[228] Before killing him, Dahmer decided he needed more beer before he could go through with killing and disposal, and he left the young man in his apartment.[229] As he returned to his apartment after having a drink, he saw Sinthasomphone sitting nude on the curb.[230] Dahmer was taking him back to his apartment when both the police and fire departments showed up.[231] Dahmer reports that he told the police that his friend always acted like this when he got drunk and did not speak English.[232] The police officers who spoke with Dahmer during this incident testified that he "responded in a calm, clear voice."[233] Dahmer related that the youth was his friend who had drunk too much and passed out on the couch.[234] Dahmer spoke coherently and did not appear to be drunk.[235] He spoke with the officers about how bad crime was in the neighborhood, and brought them into his apartment they described as "well-kept" and "neat."[236] The officers found Polaroid photos of Sinthasomphone in the apartment, which they saw as confirming Dahmer's story. They left Dahmer with what soon became his thirteenth victim. Dahmer would later state that after the police left, he gave Sinthasomphone another shot of muriatic acid and killed him.[237] Furthermore,

223. David Doege, *Those Who Got Away Recall Dahmer 'Why Is He Looking at Me Like That?' Victim Asked*, MILWAUKEE SENTINEL, Feb. 8, 1992, at 1.
224. *Id.*
225. *Id.*
226. *Id.*
227. *Id.* (alteration in original) (internal quotation marks omitted).
228. Fosdal, *supra* note 34, at 45.
229. *See* Dennis Murphy, Statement of Jeffrey Dahmer, Case No. 2472 § 5, at 117 (Aug. 2, 1991) (unpublished police report, on file with author).
230. *Id.*
231. *Id.*
232. *Id.*
233. David Doege, *Officers Recall Dahmer, Youth*, MILWAUKEE SENTINEL, Feb. 12, 1992, at 1 (internal quotation marks omitted).
234. *Id.*
235. *Id.*
236. *Id.* (internal quotation marks omitted).
237. Fosdal, *supra* note 34, at 46.

Dahmer later told police that the body of another victim, Tony Hughes, was on the floor of his bedroom at that time.[238] This evidence underscored Dahmer's control of an extraordinarily stressful situation within an hour of committing a murder. It undermines the defense's claim that he was beset by unchecked passion.

The prosecution's expert witnesses largely corroborated the impressions of the lay witnesses that Dahmer could control his actions at all times. Echoing Freud's observation quoted above, Dr. Frederick Fosdal testified that he had never seen a sexual disorder that rendered someone unable to follow the law.[239] Although Dahmer enjoyed the sex with a compliant partner, he did not enjoy the killing,[240] and actually let some folks go because he didn't have the energy to kill when he had a hangover.[241] He found killing difficult unless he was somewhat drunk.[242] Under Fosdal's analysis, the desire for sex was separated from the unpleasant task of killing and the administrative details of disposing of the corpses.[243] Fosdal further undercut Dahmer's claim of uncontrollable passion by eliciting that Dahmer would only approach men who did not have cars so they would not leave evidence outside his apartment.[244]

The theme of the final prosecution witness, Dr. Park Elliott Dietz, was that none of Dahmer's acts were impulsive. Rather, each charged

---

238. *See* David E. Umhoefer, *Police Were in Dahmer Flat*, MILWAUKEE J., Aug. 1, 1991, at 1.

239. Jim Stingl, *Disorder Alone Is Not Insanity, Expert Testifies*, MILWAUKEE J., Feb. 10, 1992, at 1.

240. *Testimony from the Dahmer Trial*, MILWAUKEE SENTINEL, Feb. 11, 1992, at 8:

> E. Michael McCann, district attorney: Did the defendant at any time say to you that he enjoyed the killing?
> Fosdal: Repeatedly he denied that. I think that is established.
>
> McCann: Did he say anything about taking pleasure from the killing?
> Fosdal: He repeatedly denied that.
> . . . .
> McCann: So the pleasure was the sex before and after, but he did not have that powerful . . . motive, a desire for the killing?
> Fosdal: That was an unwanted step.

(second alteration in original).

241. Fosdal, *supra* note 28, at 25.

242. *Testimony from the Dahmer Trial*, *supra* note 240, at 8.

243. *Id.*

244. Fosdal, *supra* note 30, at 59:

> Q: You and a guy go home together in a car—[h]e gives you a ride home and then he's done with and then his car is parked down in the street.
> A: That wouldn't have worked.
>
> Q: Was that an issue?
> A: Yeah—if they had a car, then I wouldn't ask them back.
>
> Q: You meet the guy at the tavern, and he says I have a car.
> A: Then I wouldn't have pursued it any further.
> . . . .
> A: They would have parked the car near the house and that wouldn't have worked—[t]hey could have been traced.

killing was "a planned and deliberate act."[245] Dahmer would grind sleeping pills before he went out to find a victim so they would be ready to mix in a drink.[246] He was able to be charming, seductive, and lure people back to his apartment. He would kill only on weekends so he could spend more time with the bodies and not have to go to work.[247] He only killed in his own apartment where he could control who could come in so that he would not be bothered.[248] Dahmer also related to Dietz that he knew right from wrong every time he killed, and he could have stopped himself from killing had someone walked in on him just before he committed the act.[249] Dahmer also said that if he could have "obtained the company of these men and had sexual contact with them with less drastic means, he would have stopped" killing.[250]

Perhaps the most telling testimony was Dietz's report that Dahmer explained he had always used a condom when engaging in sex with a corpse or unconscious person to avoid contracting AIDS or other diseases.[251] "The intensity of his sexual urge at that point was less than that many teenagers experience in the back seat with their girlfriend," Dietz testified.[252] This observation destroyed Dahmer's claim of unbridled and uncontrollable passion.

Dietz also drew out that necrophilia was not Dahmer's primary attraction.[253] Dahmer's first desire was for an attractive sexual partner who would be under his complete control and never leave him.[254] His preference would have been for an enduring relationship with an attractive living person, but he never found someone who fit these criteria, so he "settle[d] for less attractive, paraphilic alternatives."[255] His second choice would be a "zombie" sexual partner who would be alive indefinitely but would be lacking in will and therefore submit to his wishes.[256] Dietz observed that this fantasy is not uncommon and is played out in horror films with science fiction plots such as *The Stepford Wives*.[257] Dahmer's next choice would be an unconscious sexual partner.[258] This sort of object also appears in Western cultures in the fairy tale *Sleeping Beauty*.[259] Only if these choices were unavailable would Dahmer begin to fantasize

245. Jim Stingl, *Dahmer Knew What He Was Doing, Expert Says*, MILWALKEE J., Feb. 12, 1992, at 1 (internal quotation marks omitted).
246. *Id.*
247. *Id.*
248. *Id.*
249. *Id.*
250. *Id.* (internal quotation marks omitted).
251. Jim Stingl, *Prosecution Expert Doesn't Budge*, MILWAUKEE J., Feb. 13, 1992, at 1.
252. *Id.* (internal quotation marks omitted).
253. Dietz, *supra* note 23, at 2.
254. *Id.*
255. *Id.*
256. *Id.*
257. *Id.*
258. *Id.*
259. *Id.*

about a freeze-dried body or freshly dead corpse of an attractive man.[260] Despite these unorthodox sexual longings, Dietz concluded that Dahmer was able to function well in society and did not meet the criteria of any of the recognized patterns of personality disorder.[261]

CONCLUSION

Given the jury's task of constructing a master narrative, weaving together the disparate strands of testimony to create a coherent picture, it is perhaps understandable that the jury rejected Dahmer's claim of insanity.[262] There were too many logical gaps in a context too unfinished to account for all the evidence of his complex life. The defense did not give the jury a way of connecting the uncontrollable necrophile they portrayed with a person who functioned rationally both at work and in his other human interactions. The defense's case seemed pretextual because the facts they elicited appeared incomplete in light of the broader context revealed by Dahmer's own statements and actions.

Ultimately, Dahmer's silence, which made his killings possible, may have led to his downfall in court. Soren Kierkegaard's *Fear and Trembling* begins with a series of meditations on the Biblical patriarch Abraham's failure to speak as he was taking his son Isaac up Mount Moriah.[263] By remaining silent, by not informing Isaac that he was told to sacrifice him, Kierkegaard maintains Abraham failed to act ethically.[264] Derrida observes in his commentary on Kierkegaard, "[Abraham] speaks and doesn't speak. . . . He speaks in order not to say anything about the essential thing that he must keep secret."[265]

Isn't this what Dahmer does both on the street and in the courtroom? He does not tell those whom he is seducing that they are potential prey, that he is willing to kill them if they refuse to follow his every whim. By strangling them, he silences their voices so they cannot be witnesses against him. He does not respond to "missing" advertisements he sees in the papers. Then, in the courtroom, he mutes his own voice so the jury cannot observe his self-interested and manipulative behavior first hand; rather, he is audible only in the voices to which he has chosen to describe his past, a narrative that he may well have constructed for his own purposes. While clothing himself in the guise of a Romeo or Juliet,

260. *Id.*
261. *Id.* at 6.
262. RICOEUR, *supra* note 163, at 178.
263. SOREN KIERKEGAARD, FEAR AND TREMBLING AND THE SICKNESS UNTO DEATH 26–77 (Walter Lowrie trans., Princeton Univ. Press 3d ed. 1968) (1843 & 1849).
264. *Id.* at 67; DERRIDA, *supra* note 2, at 60.
265. DERRIDA, *supra* note 2, at 60.

a lover willing to give of self for the other, he in fact used the other for his own selfish purposes.[266]

Ultimately, the weight of the testimony—the text generated at trial—persuaded the jury that the silent defendant's proposed history of events was incomplete. The interpretation of the past he proffered was riddled with gaps and inconsistencies that could not be reconciled with his actions, his previous statements, and the observations of others. Unlike his trusting victims, the jury refused to be moved by the defendant's silence. They found the context proposed by the prosecution more compelling, and rather than accepting the defendant's muteness, they spoke in his place.

266. *See* Stingl, *supra* note 213, at 1 (according to Detective Murphy, Dahmer claimed he killed "[f]or his own warped selfish desire for self-gratification" (internal quotation marks omitted)).

CERTIFICATION AFTER *ARIZONANS FOR OFFICIAL ENGLISH V. ARIZONA*: A SURVEY OF FEDERAL APPELLATE COURTS' PRACTICES

MOLLY THOMAS-JENSEN†

INTRODUCTION

Certification of state law questions to state courts allows a dialogue between federal and state courts over questions of common concern. The process of certification generally involves a federal court sending a difficult or novel question of state law to that state's highest court, whereupon the state court responds with an answer that is authoritative. It does not, however, require the federal court to relinquish jurisdiction over the proceedings while the parties litigate the issue in state court. As a result, certification has the potential to diminish tension between federal and state courts' respective roles in a federalist system while improving efficiency and accessibility for litigants.

While certification is not without its critics, it has been an important component of the federal courts' toolkit since the Supreme Court initially endorsed the use of certification in *Clay v. Sun Insurance Office Ltd.*[1] In the years since *Clay*, more states have adopted certification procedures and more federal courts have relied upon certification when faced with novel questions of state law. In *Arizonans for Official English v. Arizona*,[2] the Supreme Court expanded the role of certification when it directed lower courts to certify questions of state law in cases where previously they would have abstained under the *Pullman* doctrine—a doctrine which allows federal courts to abstain in certain cases involving federal constitutional challenges to state laws.[3]

In this article, I examine the development of certification law in the years following the Supreme Court's decision in *Arizonans for Official English*. This article seeks to determine whether federal appellate courts have followed the Supreme Court's directive to certify when previously they would have abstained pursuant to the *Pullman* doctrine. Specifical-

---

† A.B., Brown University; J.D., Harvard Law School. I am grateful for the comments and suggestions on earlier drafts provided by: Rishi Batra, Dario Borghesan, Richard Fallon, Alison Kamhi, Mona Lewandoski, and Deborah Popowski.

1. 363 U.S. 207 (1960). *Clay* was a federal diversity case that involved a novel question of Florida law. The Court commended the Florida Legislature for its "rare foresight" in passing a law that allowed federal courts to certify questions of state law to the Florida Supreme Court. *Id.* at 212.

2. 520 U.S. 43 (1997).

3. *Id.* at 75–76. The Supreme Court first announced the *Pullman* doctrine in *Railroad Commission of Texas v. Pullman Co.*, 312 U.S. 496 (1941). The doctrine is discussed more fully in Part II of this article.

ly, this article catalogues the analytical frameworks federal courts use when determining whether to certify a question of state law in a case that presents *Pullman* considerations and assesses whether these frameworks are consistent with *Arizonans*—and, as an intertwined inquiry, whether they produce *results* that are consistent with *Arizonans*. After concluding that the federal appellate courts' analytical approach to certification in *Pullman*-type cases is inconsistent and frequently at odds with *Arizonans*, this article recommends that courts follow a formal framework that promises analytical consistency and compliance with the directive of *Arizonans*.

Since *Arizonans for Official English*, the Supreme Court has remained relatively silent on questions concerning when and whether to certify. The *Arizonans* decision provided a relatively clear directive—certify when previously you would have abstained pursuant to *Pullman*—but the federal appellate courts have not implemented this rule consistently or predictably. Both *Pullman* abstention and *Arizonans* certification promote federalism values such as comity,[4] but *Arizonans* certification nearly always improves judicial efficiency while reducing the parties' litigation time and costs.[5] By adopting a clear and consistent framework to analyze whether certification is appropriate in cases involving *Pullman* considerations, the courts of appeals will provide valuable guidance to federal district courts. They will also promote the values that are at the core of the *Arizonans* and *Pullman* decisions—namely, federal-state comity and the efficient use of both the judiciary's and litigants' resources.

To provide a foundation for the discussion of the federal appellate courts' implementation of *Arizonans for Official English*, Part I begins with a description of the formation of the *Pullman* doctrine as well as the *Arizonans* decision and its impact on the doctrine of *Pullman* abstention. Part II discusses the doctrinal and pragmatic considerations that might lead a court to find abstention preferable to certification, even after the Supreme Court's decision in *Arizonans*. I conclude that there are relatively few considerations that should sway a court to abstain rather than certify, but note that these considerations remain important and should inform courts' analyses. With this framework in place, Part III provides an overview of the federal appellate courts' approaches to analyzing whether to abstain or certify in *Pullman*-type cases decided since *Arizonans*. The analysis has varied dramatically, from circuit to circuit, and even within circuits. I conclude that the vast range of analytical approaches used by the federal appellate courts produces confusing case

4. *See generally* Guido Calabresi, *Federal and State Courts: Restoring a Workable Balance*, 78 N.Y.U. L. REV. 1293 (2003) (describing a "crisis" in the balance between state courts and lower federal courts and prescribing more federal court certification as one way to address the crisis in comity).

5. *See infra* Part I.C.

law that is often at odds with *Arizonans*. In Part IV, I propose an analytic framework that has the potential to provide clarity and consistency to this rather muddled field of law.

I. RAILROAD PORTERS AND THE LANGUAGE OF STATE GOVERNMENT: A SUMMARY OF THE *PULLMAN* AND *ARIZONANS FOR OFFICIAL ENGLISH* DECISIONS

A. Railroad Commission of Texas v. Pullman Co.[6]

In 1941, the Supreme Court decided *Railroad Commission of Texas v. Pullman Co.* The case involved a challenge to the Railroad Commission's decision to require that all sleeping cars operated on Texas rail lines be continuously attended by Pullman conductors.[7] It had been the Pullman Company's practice, on lesser-traveled routes with only one sleeper car on a train, to staff the train with a Pullman porter rather than a conductor.[8] As the United States Supreme Court explained, it was "well known" that Pullman porters were African-American while Pullman conductors were white.[9] The Pullman porters intervened in the lawsuit brought by the Pullman Company challenging the Commission's order, arguing that the order was "a discrimination against Negroes in violation of the Fourteenth Amendment."[10] The cases also presented an unsettled issue of state law, namely whether the Railroad Commission had the power to issue the order that the Pullman Company and Pullman porters were challenging.[11]

Rather than decide the constitutional issue,[12] the Court turned instead to the state issue.[13] The Court held that the federal district court, when presented with the question of state law, should have "exercise[d] its wise discretion by staying its hands."[14] In explaining its decision, the Court emphasized the "sensitive" nature of the constitutional issue and determined that "[s]uch constitutional adjudication plainly can be avoided if a definitive ruling on the state issue would terminate the controversy."[15] In a nutshell, the *Pullman* doctrine today stands for the proposition that a federal court, when faced with a tough constitutional question that could be avoided if a state law question were decided in a certain way, should exercise its discretion to abstain from addressing the state law question if it is difficult, sensitive, or relates to an important

6. 312 U.S. 496 (1941).
7. *Id.* at 497–98.
8. *Id.* at 497.
9. *Id.*
10. *Id.* at 498.
11. *Id.* at 498–99.
12. Note that the Court heard this case over a decade before deciding *Brown v. Board of Education*, 347 U.S. 483 (1954).
13. *See Pullman*, 312 U.S. at 501.
14. *Id.*
15. *Id.* at 498.

governmental function.[16] The federal court should then direct litigants to bring their claims in state court, so that a state court may decide novel or unsettled questions of state law. If the federal issue is not precluded by the decision in the state court, the litigants may return to federal court for the resolution of the case.[17]

The doctrine of *Pullman* abstention remains an important limitation on the role of federal courts. The principles underlying the Supreme Court's directive to abstain from answering the state law question in *Pullman* remain vital to ongoing discussions about the proper role of federal courts, notably in that the opinion's reasoning relied upon the Court's assessment that abstention could promote federal–state comity. The *Pullman* Court emphasized that, at least in cases brought at equity, an "unnecessary ruling" on a question of state law made by a federal court would not further the "reign of law" if it were subsequently supplanted by a state court decision.[18] The Court also reasoned that such a ruling might create friction between the federal and state systems.[19]

This doctrine is not without limitations. The most relevant to this article is a policy against abstaining in cases in which plaintiffs seek to vindicate First Amendment rights. The Supreme Court has explained that, "[i]n such case to force the plaintiff who has commenced a federal action to suffer the delay of state court proceedings might itself effect the impermissible chilling of the very constitutional right he seeks to protect."[20] Federal appellate courts have generally followed this rule, and have declined to abstain in cases involving allegations that a state law violates the First Amendment.[21]

---

16. *See* Arizonans for Official English v. Arizona, 520 U.S. 43, 76 (1997):

> Designed to avoid federal-court error in deciding state-law questions antecedent to federal constitutional issues, the *Pullman* mechanism remitted parties to the state courts for adjudication of the unsettled state-law issues. If settlement of the state-law question did not prove dispositive of the case, the parties could return to the federal court for decision of the federal issues.

For the purposes of this article, I refer to cases that present these characteristics—a federal challenge to a state law that has not yet been interpreted by that state's highest court—as a case presenting or involving "*Pullman* considerations." This simply means that these are cases in which a court prior to *Arizonans* would have looked to the *Pullman* decision and weighed these factors before determining, in its discretion, whether abstention was proper.

17. *See* England v. La. State Bd. of Med. Exam'rs, 375 U.S. 411, 421–22 (1964).

18. *Pullman*, 312 U.S. at 500.

19. *Id.*

20. Zwickler v. Koota, 389 U.S. 241, 252 (1967); *see also* Baggett v. Bullitt, 377 U.S. 360, 378–79 (1964) ("We also cannot ignore that abstention operates to require piecemeal adjudication in many courts, . . . thereby delaying ultimate adjudication on the merits for an undue length of time, . . . a result quite costly where the vagueness of a state statute may inhibit the exercise of First Amendment freedoms." (citations omitted)).

21. *See, e.g.*, Bad Frog Brewery, Inc. v. N.Y. State Liquor Auth., 134 F.3d 87, 94 (2d Cir. 1998) ("If abstention is normally unwarranted where an allegedly overbroad state statute, challenged facially, will inhibit allegedly protected speech, it is even less appropriate here, where such speech has been specifically prohibited. Abstention would risk substantial delay while Bad Frog litigated its state law issues in the state courts."); Cate v. Oldham, 707 F.2d 1176, 1184 (11th Cir. 1983) ("Ab-

Subsequent cases have determined *Pullman*'s applicability to new situations, including claims that invoke federal statutes rather than the federal constitution,[22] civil rights claims brought under 42 U.S.C. § 1983,[23] instances in which the federal court abstains and the state court decides both the state and the federal issues,[24] and cases involving any type of discretionary relief.[25] In the process, the Supreme Court has ensured the longevity and continued relevance of the *Pullman* doctrine.

B. Arizonans for Official English v. Arizona[26]

In 1997, the Supreme Court decided *Arizonans for Official English v. Arizona*. The Court relied upon a line of cases supporting the federal courts' use of state law certification procedures in holding that the federal court should have certified the central state law issue.[27] This holding established a new set of rules for courts faced with cases that involve *Pullman* considerations.

Arizonans involved a federal constitutional challenge to an amendment to the Arizona Constitution that declared English to be Arizona's official language and "the language of . . . all government functions and actions."[28] A state employee, Maria-Kelly Yniguez, working as an insurance claims manager for the state's Department of Administration brought suit in federal court against the State of Arizona, Arizona's governor, and several other state officials.[29] She claimed that the Arizona amendment violated the First and Fourteenth Amendments to the U.S.

---

stention is to be invoked particularly sparingly in actions involving alleged deprivations of First Amendment rights . . . .").

22. *See, e.g.*, Propper v. Clark, 337 U.S. 472, 490, 492 (1949) (holding that *Pullman* abstention was not appropriate to avoid the decision of sensitive non-constitutional issues).

23. *See, e.g.*, Harrison v. NAACP, 360 U.S. 167, 169, 176–77 (1959) (holding that *Pullman* abstention is appropriate in cases brought under 42 U.S.C. §1983). Section 1983 provides a cause of action for many plaintiffs alleging federal constitutional violations. It provides, in relevant part:

> Every person who, under color of any statute, ordinance, regulation, custom, or usage, of any State or Territory or the District of Columbia, subjects, or causes to be subjected, any citizen of the United States or other person within the jurisdiction thereof to the deprivation of any rights, privileges, or immunities secured by the Constitution and laws, shall be liable to the party injured in an action at law, suit in equity, or other proper proceeding for redress . . . .

42 U.S.C. § 1983 (2006).

24. *See, e.g.*, England v. La. State Bd. of Med. Exam'rs, 375 U.S. 411, 421 (1964) (holding that, when the state court addresses both state and federal issues, the federal plaintiff is bound by res judicata only if the plaintiff "voluntarily . . . and fully litigated his federal claims in the state courts").

25. Quackenbush v. Allstate Ins. Co., 517 U.S. 706, 719, 731 (1996) (discussing general principles of abstention and noting "we have recognized that the authority of a federal court to abstain from exercising its jurisdiction extends to all cases in which the court has discretion to grant or deny relief").

26. 520 U.S. 43 (1997).

27. *Id.* at 77 (citing Virginia v. Am. Booksellers Ass'n, Inc., 484 U.S 383, 393–96 (1988); Bellotti v. Baird, 428 U.S. 132, 148 (1976); Lehman Bros. v. Schein, 416 U.S. 386, 391 (1974)).

28. *Id.* at 48 (alteration in original) (quoting ARIZ. CONST. art. XXVIII, §1(1)–(2)).

29. *Id.* at 48–50.

Constitution.[30] Yniguez, who was fluent in both Spanish and English, alleged that her work handling medical malpractice claims required significant interactions with the public and that, when working with members of the public who only spoke Spanish, she spoke with them in Spanish.[31] She brought suit because she believed that her Spanish-language interactions might lead to the termination of her employment or other sanctions, and she requested injunctive and declaratory relief that the Arizona amendment violated the U.S. Constitution.[32]

Prior to trial, the Arizona Attorney General issued an official opinion that the amendment to the Arizona Constitution must be read to apply only to "official acts of government," and therefore that it did not apply to "the delivery of governmental services."[33] The attorney general reached this conclusion, at least in part, because he was obliged to read this amendment "in line . . . 'with the United States Constitution.'"[34] The Arizona Attorney General had asked both the district court and the court of appeals "to seek, through [Arizona's] certification process, an authoritative construction of the new measure from the Arizona Supreme Court."[35] Despite the Arizona Attorney General's opinion and request for certification, both the United States District Court for the District of Arizona and the Ninth Circuit Court of Appeals declined to certify the issue to the Arizona Supreme Court and found that the amendment was "fatally overbroad."[36]

The U.S. Supreme Court chastised the lower courts for failing to certify to the Arizona Supreme Court the issue of the proper construction of the constitutional amendment. The Court declared that "[c]ertification today covers territory once dominated by a deferral device called '*Pullman* abstention.'"[37] The Court emphasized that federal courts should not reach constitutional questions unless necessary to resolve the case. The Court reasoned that the Arizona Attorney General's actions in this case—namely, issuing an opinion that indicated Yniguez's actions were not unlawful under the Arizona amendment and requesting that the district court and the court of appeals certify the issue to the Arizona Supreme Court—suggested that the constitutional issue might be easily avoided by a narrowing construction of the Arizona amendment.[38] Indeed, the Court's opinion notes that the amendment's sponsors affirmed at oral argument before the Court that the attorney general's narrow construction was the correct one—and this construction, all parties seemed

30. *Id.* at 50.
31. *Id.*
32. *Id.* at 50–51.
33. *Id.* at 52 (quoting Op. Att'y Gen. No. I89-009 (1989)).
34. *Id.* (quoting Op. Att'y Gen. No. I89-009 (1989)).
35. *Id.* at 75.
36. *Id.* at 54–55, 62.
37. *Id.* at 75.
38. *Id.* at 75, 77.

to agree, would avoid the constitutional issue presented by this case.[39] Finally, the Court noted that this question was particularly appropriate for the state courts to decide because it touched upon a matter of "importance to the conduct of Arizona's business."[40]

In determining that the lower federal courts should have certified the state law issue to the Arizona Supreme Court, the *Arizonans* Court noted that *Pullman* abstention was "[a]ttractive in theory" but "protracted and expensive in practice."[41] The Court emphasized, by comparison, the virtues of certification, inasmuch as it "allows a federal court faced with a novel state-law question to put the question directly to the State's highest court, reducing the delay, cutting the cost, and increasing the assurance of gaining an authoritative response."[42] The Court relied upon previous cases in which it had found certification the appropriate course of action[43]—most notably, *Bellotti v. Baird*,[44] which held certification to be appropriate in a federal constitutional challenge to a state law restricting access to abortion.[45] Today, *Arizonans for Official English* stands for the idea that federal courts should certify where previously they would have abstained under the *Pullman* doctrine.[46]

C. The Relative Merits of Certification

As *Arizonans for Official English* emphasized,[47] *Pullman* abstention involves considerable expense to litigants, lengthens the litigation timeline, and makes inefficient use of judicial resources.[48] In a 1977 law review article, Martha Field summarized the problems associated with abstention:

> [E]xtreme delays inherent in the abstention procedure, and the attendant expense, have been chronicled many times. Essentially, the parties to a case in which abstention is ordered must undergo two law-

---

39. *Id.* at 78.

40. *Id.*

41. *Id.* at 76.

42. *Id.* Note that certification is not limited to *Pullman*-type situations. Indeed, certification may be an appropriate option in federal cases not invoking federal question jurisdiction. *See* RICHARD H. FALLON, JR., DANIEL J. MELTZER & DAVID L. SHAPIRO, HART AND WECHSLER'S THE FEDERAL COURTS AND THE FEDERAL SYSTEM 1201 (5th ed. 2003) (describing cases prior to *Arizonans for Official English*, in which the Supreme Court endorsed the use of certification in cases not involving federal constitutional challenges to state laws).

43. *See Arizonans for Official English*, 520 U.S. at 76.

44. 428 U.S. 132 (1976).

45. *Id.* at 151:

> The importance of speed in resolution of the instant litigation is manifest. Each day the statute is in effect, irretrievable events, with substantial personal consequences, occur. Although we do not mean to intimate that abstention would be improper in this case were certification not possible, the availability of certification greatly simplifies the analysis.

46. *See Arizonans for Official English*, 520 U.S. at 75. Note, however, that *Arizonans* had other holdings unrelated to the abstention/certification issue. *See id.* at 64–67 (standing); *id.* at 68–71 (mootness).

47. *Id.* at 77.

48. *See* Martha A. Field, *The Abstention Doctrine Today*, 125 U. PA. L. REV. 590, 591 (1977).

> suits instead of one, because their cause is bifurcated between state and federal courts. When a federal court abstains in an action before it, the plaintiffs must commence a new lawsuit in state trial court, usually a declaratory judgment action, to have the unclear issue of state law resolved. They must work their way up through the state appellate system, usually without getting any priority on crowded state dockets, before the state issue is settled, so that they can return to the federal system for resolution of federal issues, with the attendant appeals. The prospect is hardly a happy one for litigants. It may deter them from seeking a federal forum in the first instance, or it may, once abstention is ordered, induce them to cut their costs by presenting all issues to the state court for decision and waiving their right to return to federal court on the federal issues.[49]

While certification is generally more efficient and entails less delay than abstention, the procedure is not without its critics.[50] Nevertheless, faced with a choice between certification and abstention, certification will almost always be the speedier, more efficient, and most cost-effective option.[51]

Some who criticize certification (and abstention, for that matter) have questioned whether these procedures actually promote federal–state comity, as their supporters claim.[52] For instance, in an article criticizing what he viewed as the overuse and misuse of certification, Judge Selya of the First Circuit Court of Appeals wrote, "In the end, I believe that it engenders more understanding, and a healthier respect for state courts and what they do, when federal courts tackle the complexities of state law head on."[53] Others, such as Jonathan Remy Nash, question whether certification is consistent with the constitutional and statutory limits on

---

49. *Id.*

50. *See, e.g.*, Bruce M. Selya, *Certified Madness: Ask a Silly Question . . .*, 29 SUFFOLK U. L. REV. 677 *passim* (1995). *But see* Eric Eisenberg, *A Divine Comity: Certification (at Last) in North Carolina*, 58 DUKE L.J. 69, 77–81 (2008) (noting that a well-drafted certification procedure, combined with the state's highest court's judicious use of its discretion, could avoid many of the pitfalls of certification).

51. *See, e.g.*, Rebecca A. Cochran, *Federal Court Certification of Questions of State Law to State Courts: A Theoretical and Empirical Study*, 29 J. LEGIS. 157, 217 (2003) (cataloguing delays of over one year in Ohio's response to certified questions, but noting that "even waiting a year for an answer to a certified question answer pales when compared to the time elapsed under abstention"); Judith S. Kaye & Kenneth I. Weissman, *Interactive Judicial Federalism: Certified Questions in New York*, 69 FORDHAM L. REV. 373, 397 (2000) (noting that New York courts have generally responded to certified questions of law within six months of accepting the certified question); *see also* Jonathan Remy Nash, *Examining the Power of Federal Courts to Certify State Law*, 88 CORNELL L. REV. 1672, 1698 (2003) ("[C]ertification avoids procedural complications that might hinder the state court system's resolution of the state law question were abstention employed."). *But see* Selya, *supra* note 50, at 688 (noting that both certification and abstention require "piecemeal litigation spanning two separate court systems—and such divisions are notoriously inefficient").

52. *See, e.g.*, Selya, *supra* note 50, at 684–87 (emphasizing that certification is a one-way street and noting that to allow state courts to certify questions of federal law to federal courts would undermine the constitutional role of state courts in the federal system).

53. *See id.* at 687.

federal jurisdiction.[54] Nash has argued that certification "raises serious questions of federal jurisdiction," and that its "jurisdictional underpinnings" are inherently weak.[55]

Additionally, some judges and practitioners have been critical of the results that certification produces in practice. Several commentators have noted that problems arise when federal courts do not "artfully present[]" the questions that they wish answered.[56] Other commentators have criticized state courts when they dismiss a certified question without explanation, arguing that "a brief, reasoned opinion when declining to answer would help avoid federal court frustration and misinterpretation of silence, as well as expressly enforce the policies and practice of [state] certification."[57] Judge Selya has argued that "the advisory nature of certified questions poses a psychological barrier to enthusiastic engagement with the issues."[58] State courts are not given jurisdiction over the cases in which questions are certified to them,[59] but their answers are binding on the federal court and on future litigants.[60] Judge Selya concludes that this lack of jurisdiction and the resulting "psychological barrier" leads state courts to treat the certified issue differently than they would if it were a live case or controversy that had worked its way through the state court system.[61]

Despite these criticisms of both the theory and practical problems underlying certification, it is generally accepted that certification is preferable to the lengthy and expensive process associated with abstention.[62]

---

54. Nash, *supra* note 51, at 1672 (concluding that certification can be understood under both a "unitary" and a "binary" conception, but that neither framework can provide a full foundation for both constitutional and statutory jurisdiction).

55. *See id.* at 1748–49.

56. Selya, *supra* note 50, at 689; *see also* Cochran, *supra* note 51, at 198 (noting that most questions certified to the Ohio Supreme Court failed to meet the court's requirement that the certified question be a pure question of law); Kaye & Weissman, *supra* note 51, at 404–414 (outlining cases in which the New York Court of Appeals declined to answer certified questions).

57. Cochran, *supra* note 51, at 179; *see also* Selya, *supra* note 50, at 681–82; Richard Alan Chase, Note, *A State Court's Refusal to Answer Certified Questions: Are Inferences Permitted?*, 66 ST. JOHN'S L. REV. 407, 422 (1992).

58. Selya, *supra* note 50, at 682.

59. *But see* Nash, *supra* note 51, at 1675:

> First, the unitary conception applies when a federal court that certifies a question to a state high court is understood to transfer the very case that is before it (or some portion thereof) to the state court. The state court considers and responds to the questions of state law, whereupon the federal court regains control of the case. Under this conception, there is one, unitary case, with jurisdiction shifting from the federal court to the state court and then back to the federal court.

60. 17A CHARLES ALAN WRIGHT ET AL., FEDERAL PRACTICE AND PROCEDURE: JURISDICTION 3d §4248 (2007 & Supp. 2009).

61. Selya, *supra* note 50, at 682.

62. *See, e.g.*, *id.* at 688 (arguing that both certification and abstention are problematic and that federal courts should "abjur[e]" from piecemeal litigation, but not disputing the claim that certification, as compared to abstention, is somewhat more efficient); *see also* Kaye & Weissman, *supra* note 51, at 384–85:

> Supreme Court support for certification has been widely echoed. In 1967, the National Conference of Commissioners on Uniform State Laws adopted the Uniform Certification

And, given the Supreme Court's strong endorsement of the use of certification in *Pullman*-type cases,[63] it appears that, for these cases at least, certification is here to stay.

II. SITUATIONS IN WHICH A COURT MIGHT CHOOSE TO ABSTAIN RATHER THAN CERTIFY

The consensus among commentators appears to be that, when choosing between abstention and certification, federal courts ought to choose the latter, both for doctrinal and pragmatic reasons.[64] Why, then, would any court abstain, given the very tangible advantages of certification and the U.S. Supreme Court's endorsement of the certification process? This section details those scenarios in which a court may decide to abstain without contravening the Supreme Court's directive in *Arizonans for Official English*. These procedures are necessarily intertwined: as more courts certify in *Pullman*-type cases, it follows that courts should resort to *Pullman* abstention with decreasing frequency.

Federal courts' doctrine provides several scenarios in which a federal court should abstain rather than certify. In *Pennhurst State School & Hospital v. Halderman*,[65] the Supreme Court expanded its Eleventh Amendment jurisprudence to hold that a federal court could not enjoin a state official from violating state law.[66] If a federal court certified a question of state law to the state's highest court in order to avoid a federal constitutional issue under *Arizonans* and *Pullman*, the *Pennhurst* doctrine would nevertheless prevent the federal court from granting injunctive relief if the state court's response made clear that a state official's actions were not consistent with state law. In that scenario, the court would not need to reach the federal constitutional issue, because of the state court's response; but it would also be unable to grant an injunction ordering the state official to comply with state law. In other words, certification for *Pullman* reasons cannot deal with the *Pennhurst* problem, and the parties would need to re-litigate the issue in state courts. In such

---

of Questions of Law Act. . . . In 1969, the American Law Institute chimed in with its support, as did the American Bar Association in 1977. In 1992, the National Conference on State–Federal Judicial Relationships suggested that certification could enhance judicial federalism, and in 1995, the Committee on Long Range Planning of the United States Judicial Conference recommended that states without certification procedures adopt them.

63. Arizonans for Official English v. Arizona, 520 U.S. 43, 75–80; *see also* Bellotti v. Baird, 428 U.S. 132, 150–51 (1976) (emphasizing the virtues of certification in a federal constitutional challenge to a state law that had not been interpreted by the state's highest court); Lehman Bros. v. Schein, 416 U.S. 386, 390–91 (1974) (emphasizing relative merits of certification when a diversity case presents an unclear issue of state law).

64. *See* sources cited *supra* note 62.

65. 465 U.S. 89 (1984). Specifically, *Pennhurst* created an exception to the rule established in *Ex parte Young*. *Id.* at 105–06. *Young* held that, even though the Eleventh Amendment barred suits against the states, plaintiffs could sue state officials seeking relief against state action. *Ex parte* Young, 209 U.S. 123, 159–60 (1908). *Pennhurst* limited *Young*'s reach, and thus limited the exception to the Eleventh Amendment's reach.

66. *Pennhurst State Sch. & Hosp.*, 465 U.S. at 117.

a situation, certification is inadequate, and the federal court would need to abstain while the state court addressed the state law issues and determined whether an injunction would be appropriate relief.

An example of this *Pullman–Pennhurst* scenario arose in *University of Utah v. Shurtleff*.[67] In *Shurtleff*, the United States District Court for the District of Utah heard a case brought by the University of Utah against the state attorney general. The University sought equitable relief and a judgment that the University's firearm policy banning concealed weapons was: (1) protected by the First and Fourteenth Amendments of the U.S. Constitution, (2) protected by a provision of the Utah Constitution that purported to give autonomy to the University, and (3) not actually in conflict with the state's concealed weapons laws.[68] The court found that the *Pennhurst* decision would prevent it from granting the requested equitable relief—enforcement of state law against a state official.[69] Ultimately, the court decided to dismiss the state law claims against Utah's attorney general because of the Eleventh Amendment protections announced in *Pennhurst*.[70] As to the remaining federal claim, the court noted that it was left with "jurisdiction over a meaty federal constitutional claim that could be rendered moot by a favorable state court decision on either of [the] state law claims."[71] Citing *Pullman* and other cases for the principle that questions of state law should be resolved before proceeding to substantial federal constitutional questions, the court determined that *Pullman* abstention was appropriate.[72] Given that the plaintiffs would need to bring the state law claims before the state court system, certification would offer none of its usual advantages. Indeed, certification in this case might have been more inefficient—inasmuch as it likely would have involved the state's highest court answering a state law question in the abstract and then later hearing an appeal of the same issues brought by the same plaintiffs.

Federal courts' doctrine creates another situation in which abstention will generally be preferable to certification, namely, when abstention is mandated under another doctrine. Most prominent amongst the other abstention doctrines is the *Younger* doctrine.[73] *Younger* requires federal courts to abstain when an injunction is sought to stay ongoing criminal proceedings.[74] Subsequent decisions extended *Younger* to cases in which declaratory judgment was sought,[75] to some cases in which the federal

67. 252 F. Supp. 2d 1264 (D. Utah 2003).
68. *Id.* at 1266–67.
69. *Id.* at 1282.
70. *Id.* at 1281–83.
71. *Id.* at 1283.
72. *Id.* at 1284–85.
73. *See* Younger v. Harris, 401 U.S. 37 (1971).
74. *Id.* at 41. Note that, unlike *Pullman*, the *Younger* doctrine of abstention is not discretionary. *Id.*
75. Samuels v. Mackell, 401 U.S. 66, 68–69 (1971).

plaintiffs brought suit before state criminal proceedings,[76] to civil enforcement cases brought by the state,[77] and even to some cases not brought by the state but which involved important state interests.[78] It would defy logic for the federal court to certify the state law question that gave rise to *Pullman* considerations if another doctrine, such as *Younger* or *Burford*[79] or *Colorado River*,[80] would require abstention.

In addition to the limitations imposed by doctrines of federal courts and federal jurisdiction, a number of pragmatic considerations might lead a federal court to abstain rather than to certify. A federal court might note that parallel proceedings in state court already exist, such that it would be inefficient and cumbersome to certify a question to the state's highest court when there is already a case working its way through the state court system that will present the same issue to the state's highest court with a full factual record and without any of the disadvantages of certification. Alternatively, a federal court might decide that questions of state law are so predominant that abstention would actually be more efficient.[81] Finally, a federal court might decide that the state law question is so fact intensive that the state court would be unable to provide a fair and reasoned answer without its own fact-finding trial procedure.

A federal court may also be limited by the certification rules of the state whose law is in question. Certification procedures vary from state to state. In California, for example, a federal district court may not certify questions of state law to the California Supreme Court; the California Rules of Court provide for certification of questions only from "the United States Supreme Court, a United States Court of Appeals, or the court of last resort of any state, territory, or commonwealth."[82] The Delaware Supreme Court Rules prohibit that state's highest court from accepting a

76. Hicks v. Miranda, 422 U.S. 332, 348–49 (1975). *But see* Doran v. Salem Inn, Inc., 422 U.S. 922, 930–31 (1975).

77. Trainor v. Hernandez, 431 U.S. 434, 444 (1977).

78. *See, e.g.*, Pennzoil Co. v. Texaco, Inc., 481 U.S. 1, 16–17 (1987).

79. Burford v. Sun Oil Co., 319 U.S. 315, 333–34 (1943) (requiring abstention where a complex state administrative scheme that is overseen by the state courts presents a better forum for adjudicating the case at issue in the federal court).

80. Colo. River Water Conservation Dist. v. United States, 424 U.S. 800, 817–18 (1976) (creating a narrow exception to the general rule that pendant proceedings in state court do not bar related proceedings in federal court).

81. In a federal question case, these additional state law issues would likely be brought in by the parties' use of supplemental jurisdiction, under 28 U.S.C. § 1367(a) (2006). *See, e.g.*, Shegog v. Bd. of Educ. of the City of Chi., 194 F.3d 836 (7th Cir. 1999). In that case, discussed in Part III.G, *infra*, the Seventh Circuit heard an appeal involving one federal claim and many state law claims. *Id.* at 837. The Seventh Circuit abstained, rather than certifying, but the court's reasoning deviated from the *Arizonans for Official English* framework, inasmuch as the court justified not certifying because it feared that this would "short-circuit" the process of decision by state courts. *Id.* at 840. If, however, the Seventh Circuit had emphasized that certification was less appropriate than abstention, because abstention might actually be more efficient for litigants and the courts, then it would have been consistent with the holding and reasoning of *Arizonans*.

82. CAL. R. CT. 8.548.

certified question where any of the material facts are in dispute.[83] And one state, North Carolina, does not provide for any certification of questions to its highest court.[84] Furthermore, many states' certification procedures give the state's highest court the discretion whether to accept the certified question.[85] Consequently, a federal court may correctly certify a question, only to have the highest state court decline to answer. In these situations, where certification is unavailable because of state certification rules or because of a state court's exercise of its discretion, a federal court considering *Pullman* abstention should either abstain or decide the state law issue on its own.

III. WHEN DO COURTS ABSTAIN AND WHEN DO THEY CERTIFY?

As the previous sections have shown, federal courts considering *Pullman* abstention *should* certify a question of state law rather than abstain. This section analyzes when they actually certify instead of abstain. Of course, a court may decide to abstain or decide the legal issue itself if it relies upon case-specific reasons—such as those discussed in the previous section—that make certification less attractive or feasible than abstention. An examination of federal case law since *Arizonans for Official English* reveals, however, that most circuits' practices do not conform neatly to these rules. Some circuits, such as the Second Circuit, have endeavored to follow the guidance provided by the U.S. Supreme Court. Other circuits, however, have seemed less concerned with the intent of *Arizonans* and have either applied an inconsistent approach to these cases or have generally preferred abstention to certification.

This section details the varied approaches taken by different circuits, with attention to the structure and reasoning employed by the various courts of appeals as they assess whether to abstain or certify. The following catalogue of cases includes descriptions of nearly every case decided since *Arizonans for Official English* in which a federal appellate court has abstained or certified a question of state law in a case involving *Pullman* considerations. The cases are organized by circuit and in chronological order within each circuit. This list does not include cases involving questions of law from states that, at the time the case was de-

83. DEL. SUP. CT. R. 41(b).

84. *See* Eisenberg, *supra* note 50, at 71.

85. *See, e.g.*, DEL. SUP. CT. R. 41(b) ("Certification will be accepted in the exercise of the discretion of the Court only where there exist important and urgent reasons for an immediate determination by this Court of the questions certified."); HAW. R. APP. 13(a):

> When a federal district or appellate court certifies to the Hawai'i Supreme Court that there is involved in any proceeding before it a question concerning the law of Hawai'i that is determinative of the cause and that there is no clear controlling precedent in the Hawai'i judicial decisions, the Hawai'i Supreme Court may answer the certified question by written opinion.

OR. R. APP. P. 12.20(3) ("The Supreme Court will consider whether to accept a question certified to it without oral or written argument from the parties unless otherwise directed by the Supreme Court."); *see also* Cochran, *supra* note 51, at 176 (noting that the Ohio Supreme Court often declines to explain its exercise of its discretion when deciding not to answer certified questions).

cided, did not have certification procedures in place. I have also omitted a handful of unpublished cases that did not add any depth or information that was not ascertainable through other cases.

A. First Circuit

The First Circuit Court of Appeals has employed several different approaches in determining whether to certify in cases involving *Pullman* considerations.[86] At times, the court has left the decision completely within the discretion of the lower court, while at other times it has directed the district court to certify, even going so far as to provide the exact question that the district court must certify. In other cases, the First Circuit has found abstention appropriate, relying on pragmatic factors such as the existence of parallel proceedings in state court.

In *Romero v. Colegio de Abogados de Puerto Rico*,[87] the First Circuit held that the district court should have stayed the case and certified a question of law to the Supreme Court of Puerto Rico.[88] The case presented a First Amendment challenge to the Puerto Rico Bar Association's use of mandatory membership dues to purchase group life insurance for members.[89] It involved *Pullman* considerations because it was unclear whether, under the state law governing the bar association, the association had the authority to purchase life insurance using member dues.[90] The First Circuit directed the district court to certify to the Supreme Court of Puerto Rico, relying upon the United States Supreme Court's opinion in *Arizonans*, and provided the specific question that the lower court should certify.[91] The court noted that certification was especially appropriate in this case because it would "recognize[]" the role that the Supreme Court of Puerto Rico played in regulating the bar.[92]

In *Laffey v. Begin*,[93] the First Circuit took another approach to the issue of whether to abstain or certify. In a First Amendment challenge to

86. For a description of *Pullman* considerations, see *supra* note 16.
87. 204 F.3d 291 (1st Cir. 2000).
88. *Id.* at 304–05.
89. *Id.* at 293:

> Romero argued that being compelled to purchase life insurance violated his First Amendment rights in two senses. First, noting that compelled membership in a bar association infringes on his freedom of association, he claimed that Puerto Rico's interests in promulgating the statutory purposes of the bar may be sufficient to overcome his interest in not being compelled to associate and pay dues, but only in support of activities germane to the bar's legitimate purposes. Mandatory purchase of life insurance as a condition of bar membership does not, he argued, meet this germaneness test. Second, he stated that he believes in a free-market economy and is opposed to government-sponsored social programs, especially for those who are not indigent. Thus, the mandatory life insurance provision is contrary to his political philosophy, and he objects to this non-incidental expenditure on ideological grounds.

90. *Id.* at 293, 305.
91. *Id.* at 305.
92. *Id.*
93. 137 F. App'x 362 (1st Cir. 2005).

Rhode Island's election contributions law, the court found that the case involved *Pullman* considerations and noted that the parties to the lawsuit had agreed that the questions of state law "should be answered as expeditiously as possible by the state supreme court."[94] The First Circuit left to the district court's discretion the decision of whether to abstain or certify, but suggested that the district court should first consider whether an appeal of the state administrative agency's decision at issue could be taken directly to the state supreme court.[95] The court continued: "If such review is not available and other conventional routes to the state court appear for any reason to be precluded, the district court may certify the appropriate state law questions to the Rhode Island Supreme Court."[96] While this section of the opinion is dicta, the court appeared to suggest that certification should be the last choice, only to be resorted to if a case could not be brought in state court. This, of course, contravenes *Arizonans for Official English*, which emphasized that courts should certify whenever possible and only resort to abstention when certification procedures were unavailable.

Finally, in *Rivera-Feliciano v. Acevedo-Vilá*,[97] the First Circuit Court of Appeals held that *Pullman* abstention was appropriate, in large part because "arguments virtually identical to those presented here had been presented to the courts of Puerto Rico even before this suit was filed, and these are now pending before the Supreme Court of Puerto Rico."[98] The court also noted that abstention was appropriate under the *Colorado River* doctrine, a doctrine that, in very limited circumstances, "permit[s] the dismissal of a federal suit due to the presence of a concurrent state proceeding for reasons of wise judicial administration."[99] The case involved a constitutional challenge to correctional procedures and involved several regulations and questions of law that the Supreme Court of Puerto Rico had not had an opportunity to decide.[100] While the parallel state proceedings and the applicability of *Colorado River* abstention made the case a strong candidate for abstention rather than certification, the opinion does not mention certification as a possibility, let alone analyze the relative merits of certification and abstention in the case.

B. Second Circuit

Shortly after *Arizonans for Official English* was decided, the Second Circuit Court of Appeals issued a pair of opinions that provided a rigorous and thorough analytical framework for deciding whether to ab-

94. *Id.* at 363.
95. *Id.* at 364.
96. *Id.*
97. 438 F.3d 50 (1st Cir. 2006).
98. *Id.* at 61.
99. Colo. River Water Conservation Dist. v. United States, 424 U.S. 800, 819 (1976).
100. *Rivera-Feliciano*, 438 F.3d at 60–62.

stain or certify. The Second Circuit's case law is clear and relatively consistent on the issue.

In *Tunick v. Safir*,[101] the Second Circuit provided a thoughtful analysis of when, in a *Pullman* case, a court should certify rather than abstain. The case involved a First Amendment challenge to state law prohibiting public nudity, brought by a photographer who wished to stage a photo shoot involving "75 to 100 nude models arranged in an abstract formation" on a city street.[102] Writing for a Second Circuit panel, Judge Calabresi noted that:

> *Arizonans* made quite clear that, in the eyes of the Supreme Court, the device of certification provides all the benefits of *Pullman* abstention (deference in a federal system to state courts on questions of state law and statutory interpretations that avoid constitutional difficulties), while reducing greatly its drawbacks (delay and cost).[103]

Judge Calabresi's opinion also noted, though, that *Arizonans* does not require a federal court to certify whenever there is a federal constitutional challenge to a state law that is unclear or has not yet been interpreted by the state's highest court. The opinion pointed to two "right to die" cases that the Supreme Court decided shortly after *Arizonans*.[104] Those two cases—*Washington v. Glucksberg*[105] and *Vacco v. Quill*[106]—involved federal constitutional challenges to state laws prohibiting physician-assisted suicide. Both cases presented *Pullman* considerations and an opportunity to certify: a federal challenge to state laws that had yet to be construed by the states' highest courts and that could be avoided by narrowing constructions.[107] In his analysis of these cases, Judge Calabresi emphasized that the U.S. Supreme Court reached the merits of the challenges "despite the concession of the parties that, under certain interpretations, the statutes would avoid constitutional challenge."[108] The opinion explained this apparent inconsistency by looking to the federalism values that underlie both certification and abstention: *Arizonans for Official English* concerned "the very manner in which Arizona was to carry out the basic functions of state governance," while the right to die cases did not implicate "core functions of state governments."[109]

Judge Calabresi concluded that *Arizonans* and the right to die cases "in no way lessen the significance of [the] *Pullman* factors" when deter-

101. 209 F.3d 67 (2d Cir. 2000).
102. *Id.* at 68.
103. *Id.* at 73.
104. *Id.* at 74.
105. 521 U.S. 702 (1997).
106. 521 U.S. 793 (1997).
107. *Tunick*, 209 F.3d at 74.
108. *Id.*
109. *Id.* at 77.

mining whether to certify.[110] Rather, *Arizonans* and the right to die cases "put a gloss" on the *Pullman* factors and point to "other factors that are relevant to the question of certification."[111] The opinion explained that a federal court considering whether to certify "in federal constitutional litigation involving state statutes" should "look first to the *Pullman* doctrine for guidance."[112] The court explained: "This is so because, although *Pullman* abstention involves problems that certification may avoid or reduce, it still remains the doctrine whose purpose is most proximate to that of certification in cases concerning the federal constitutional validity of state laws."[113] After extensive analysis, the opinion concluded:

> The composite lesson of all these cases is that there are at least six factors that must be considered in deciding whether certification is justified. They are (1) the absence of authoritative state court interpretations of the state statute, (2) the importance of the issue to the state and the likelihood that the question will recur, (3) the presence of serious constitutional difficulties that could be avoided by a possible interpretation of the statute, (4) the capacity of certification to resolve the litigation and either to render federal constitutional decisions unnecessary or to ensure that they are inescapably before the federal court, (5) the federalism implications of a decision by the federal courts and in particular whether a decision by the federal judiciary potentially interferes with core matters of state sovereignty, and (6) the effect of the delay entailed by certification on the asserted rights at issue.[114]

Based on its application of these factors to the case at hand, the Second Circuit certified a question to the New York Court of Appeals relating to the construction of the state's public nudity statute.[115] The Second Circuit determined that: (1) the New York Court of Appeals had not interpreted the statutory provision at issue;[116] (2) the constitutional issue was "grave," noting in particular, the "tortured issue of the level of protection . . . [for] artistic or expressive nudity";[117] (3) a construction of the state statute that could avoid the constitutional issue was plausible;[118] (4) certification would either completely resolve the case or resolve all issues but the constitutional issue;[119] (5) federal courts were taking a more active role than customary in reviewing city actions due to the "heavy stream of First Amendment litigation generated by New York City in recent years," and certification might relieve some of the tension

110. *Id.* at 75.
111. *Id.*
112. *Id.* at 74.
113. *Id.*
114. *Id.* at 81.
115. *Id.* at 89.
116. *Id.* at 81.
117. *Id.*
118. *Id.* at 84.
119. *Id.* at 85.

resulting from the federal courts' oversight;[120] and (6) the decision to certify would be "the least harmful alternative" in terms of delay.[121] *Tunick* remains the most thorough and rigorous analysis to date of whether a court should certify a question of state law in a case involving *Pullman* considerations.

A year later, faced with another case involving *Pullman* considerations, the Second Circuit again chose to certify the state law issues to the New York Court of Appeals.[122] In *Allstate Insurance Company v. Serio*,[123] several insurance companies brought a First Amendment challenge to a New York law limiting their ability to make referrals to insured parties.[124] The opinion, also authored by Judge Calabresi, generally followed the framework established in *Tunick*, although it did not address each of the *Tunick* factors.[125] The opinion also forged some new ground. It held that, in a *Pullman* case where the court certified a question and the state court declined to answer the certified question, "this court, once having given the state a first shot at reading its law, has full latitude to interpret the relevant state law."[126] The Second Circuit noted an additional factor weighing in favor of certification: just as in *Arizonans for Official English*, the state attorney general in *Serio* had argued a saving construction of the statute was possible.[127]

In *Nicholson v. Scoppetta*,[128] the Second Circuit heard a constitutional challenge to the City of New York's practice of removing children from the custody of a parent who had been a victim of domestic violence.[129] The court recognized family law disputes as an area of "traditional state concern," and noted that both a concern for federal-state comity, and the fact that this was an area in which the state courts were particularly knowledgeable, weighed in favor of certification.[130] The court further noted that the state law was ambiguous,[131] and that the constitutional challenges were complex but could be avoided by a particular interpretation of the state law.[132] The court relied on both *Pullman* and *Arizonans for Official English* as it reached its conclusion that certification was the appropriate course.[133] It did not rely explicitly upon the *Tu-*

120. *Id.*
121. *Id.* at 88.
122. *See* Allstate Ins. Co. v. Serio, 261 F.3d 143, 145 (2d Cir. 2001).
123. *Id.*
124. *Id.* at 144.
125. *Id.* at 149–54.
126. *Id.* at 153 n.14.
127. *Id.* at 153.
128. 344 F.3d 154 (2d Cir. 2003).
129. *Id.* at 158.
130. *Id.* at 168.
131. *Id.* at 169.
132. *Id.* at 171–76.
133. *Id.* at 167–68.

nick framework, but the analysis in this case is not at odds with that framework.

C. Third Circuit

Since *Arizonans for Official English*, the Third Circuit Court of Appeals has heard few cases involving *Pullman* abstention or requests to certify. Consequently, the Third Circuit does not have a well-developed doctrine to guide judges in deciding whether to abstain or certify.

In *Philadelphia City Council v. Schweiker*,[134] the Third Circuit reviewed a lower court's decision to abstain in a case involving a constitutional challenge to a state agency's decision to take over the Philadelphia school district.[135] The *Schweiker* court made no mention of certification as it reviewed the district court's decision to abstain.

In *Afran v. McGreevey*,[136] the Third Circuit heard a constitutional challenge to the New Jersey governor's decision to delay the effective date of his resignation.[137] The court found that *Pullman* abstention would be inappropriate, but emphasized that certification was also inappropriate for the same reason: the relative clarity of state law.[138] The court also noted that its analysis was informed by consideration of "timing, feasibility, public policy, and plaintiffs' choice of forum."[139]

D. Fourth Circuit

Like the Third Circuit, the Fourth Circuit Court of Appeals has heard few cases involving *Pullman* considerations in the years since *Arizonans for Official English*. In a 2003 case, *PSINet, Inc. v. Chapman*,[140] the Fourth Circuit heard a case challenging an amendment to a Virginia law criminalizing sale of certain sexually explicit materials to youth.[141] The statute had been litigated extensively prior to the amendment at issue in the case.[142] In the prior litigation, the U.S. Supreme Court had certified multiple questions of state law to the Supreme Court of Virginia.[143] Without citing *Pullman*, the Fourth Circuit also chose to certify several questions of state law to the Supreme Court of Virginia. This omission is troubling. While citing to *Pullman* might seem like an unnecessary formality to some, the *Pullman* decision, and the analysis it supports, should be at the foundation of any discussion about whether to certify for *Pullman* reasons. As the Second Circuit explained in *Tunick*, courts consider-

134. 40 F. App'x 672 (3d Cir. 2002).
135. *Id.* at 674.
136. 115 F. App'x 539 (3d Cir. 2004).
137. *Id.* at 540.
138. *Id.* at 542–43.
139. *Id.* at 543.
140. 317 F.3d 413 (4th Cir. 2003).
141. *Id.* at 415.
142. *Id.* at 415–18.
143. *Id.* at 416.

ing certification should look to *Pullman* for guidance because "it still remains the doctrine whose purpose is most proximate to that of certification in cases concerning the federal constitutional validity of state laws."[144] *Arizonans for Official English* did not supplant *Pullman*; instead, it merely "put a gloss" on the guidance provided by *Pullman*.[145]

Nevertheless, it does appear that the Fourth Circuit relied upon a *Pullman*-type justification in its decision to certify: "Ascertaining the scope of the [state] law's coverage and what compliance measures would preclude conviction is necessary for resolution not only of the First Amendment claim, but also for resolution of the Dormant Commerce Clause claim."[146] The Fourth Circuit cited *Arizonans* as it explained its decision to certify, and emphasized that when a federal tribunal considers invalidating a state statute that the state's highest court has yet to interpret, the federal court should tread lightly.[147]

E. Fifth Circuit

The Fifth Circuit Court of Appeals has tended to avoid certification in favor of abstention. It has also overlooked both options in some cases, even when construing state laws that involve core state functions and that are being challenged on federal constitutional grounds.

In *Nationwide Mutual Insurance Co. v. Unauthorized Practice of Law Committee*,[148] the Fifth Circuit reviewed the district court's decision to abstain under the *Pullman* doctrine.[149] The court concluded that the district court did not abuse its discretion and had not erred in abstaining.[150] As part of its analysis, the Fifth Circuit considered whether the district court had erred in not certifying the state law issue to the Supreme Court of Texas. While the Fifth Circuit cited *Arizonans* and noted that certification was likely more efficient, it nonetheless concluded that abstention was preferable to certification because other insurance companies were currently litigating the state law issues in Texas's state courts.[151] The court therefore concluded "that the Supreme Court of Texas would be better suited to answer this question with the benefit of records generated in state court by several insurance companies than it

144. Tunick v. Safir, 209 F.3d 67, 74 (2d Cir. 2000).

145. *Id.* at 75:

> Given the shared goals of *Pullman* abstention and of the device of certification, the factors counseling the former are also suggestive of when the latter is desirable. As a result, *Arizonans*, *Quill*, and *Glucksberg* in no way lessen the significance of these *Pullman* factors. They do, however, put a gloss on them, while also pointing to other factors that are relevant to the question of certification.

146. *PSINet*, 317 F.3d at 422–23.

147. *Id.* at 424.

148. 283 F.3d 650 (5th Cir. 2002).

149. *Id.* at 652.

150. *Id.* at 657.

151. *Id.* at 656–57.

would be by receiving a certified question from one insurer with a relatively limited record on appeal."[152]

In a 2005 habeas case, *Gomez v. Dretke*,[153] the Fifth Circuit abstained based on *Pullman* and other prudential considerations. The case involved a request by a prisoner to stay his habeas appeal so that he could seek remedies in state court that had recently become available.[154] The court did not discuss certification, but certification would not have addressed all of the concerns at work in the case. Most notably, certification would not address the court's desire to allow the state prisoner the opportunity to litigate his claims in the state system in light of the possibility of new relief.

In 2006, the Fifth Circuit heard *Center for Individual Freedom v. Carmouche*,[155] a First Amendment challenge to Louisiana's Campaign Finance Disclosure Act.[156] The court construed the Act "in a way that save[d] it from constitutional infirmity."[157] It did not, however, abstain or certify to allow the state courts to construe the Act. The dissent argued that "[t]he majority erred in refusing to certify the *res nova* state law questions implicated in the interpretation of the CFDA to the Louisiana Supreme Court."[158] The dissent further argued that the case involved core state functions, namely the promotion of "genuinely democratic elections to fill its major public offices free from corruption and other undue influences," and that this made certification all the more appropriate.[159]

F. Sixth Circuit

In the years immediately following *Arizonans for Official English*, the Sixth Circuit Court of Appeals tended to abstain without addressing whether certification was appropriate. While the Sixth Circuit was slow to adopt certification, it now routinely certifies questions of state law to the state court in cases involving *Pullman* considerations, although it often does so without reference to *Pullman* or discussion of abstention. For the reasons discussed above, acknowledgment of and reference to the *Pullman* doctrine provides valuable guidance to courts as they determine whether certification is the appropriate action.[160]

In 1998, the Sixth Circuit heard *Slyman v. City of Willoughby*,[161] a case involving a constitutional challenge to a zoning ordinance.[162] The

152. *Id.* at 656–57.
153. 422 F.3d 264 (5th Cir. 2005).
154. *Id.* at 265.
155. 449 F.3d 655 (5th Cir. 2006).
156. *Id.* at 658.
157. *Id.* at 664.
158. *Id.* at 672.
159. *Id.* at 669. The majority opinion does not address the dissent's arguments that abstention or certification would have been more appropriate. *See id.* at 666.
160. *See, e.g.*, Tunick v. Safir, 209 F.3d 67, 74 (2d Cir. 2000).
161. 134 F.3d 372, 1998 WL 24990 (6th Cir. 1998) (unpublished table decision).

court found that abstention under both the *Pullman* and *Thibodaux*[163] doctrines was appropriate.[164] The *Thibodaux* doctrine suggests abstention is appropriate in cases that involve particularly sensitive matters relating to "sovereign prerogative," especially where the issue somehow touches upon "the apportionment of governmental powers" between local and state-wide authorities.[165] While it did not discuss the possibility of certification, the decision to abstain rather than certify might have been due to the necessity of abstaining under the *Thibodaux* doctrine.

That same year, the Sixth Circuit decided *Gottfried v. Medical Planning Services, Inc.*,[166] a federal First Amendment challenge to a state-court-issued injunction that prohibited protesters from congregating outside of an abortion clinic.[167] The court gave a reasoned and thoughtful explanation as to the appropriateness of abstention to address the *Pullman* concerns in this case:

> Abstention is the proper course here because the state court, unlike the federal court, can modify its injunction or narrowly construe it in light of any intervening legal developments. Giving the state court the first opportunity to reassess the injunction in light of the Supreme Court's recent decisions . . . is the most efficient way to decide this case.[168]

While the opinion did not directly address whether certification was an option, this analysis of the parallel proceedings thoroughly supports the court's conclusion that abstention provided the best response to the issues present in the case.

In *Brown v. Tidwell*,[169] the Sixth Circuit upheld the district court's decision to abstain in a case involving a challenge to a state practice of collecting "jail fees."[170] The court did not discuss certification, nor did it offer any explanation that showed abstention to be particularly appropriate in the case.

In *Northland Family Planning Clinic, Inc. v. Cox*,[171] a challenge to Michigan's "partial birth abortion" law, the Sixth Circuit considered certifying a question relating to the law's meaning to the Michigan Supreme Court.[172] The case involved a federal constitutional challenge to a state law, but the Sixth Circuit held that certification—and, presumably, ab-

---

162. *Id.* at *2.
163. La. Power & Light Co. v. City of Thibodaux, 360 U.S. 25, 28 (1959).
164. *Slyman*, 1998 WL 24990, at *2.
165. *City of Thibodaux*, 360 U.S. at 28.
166. 142 F.3d 326 (6th Cir. 1998).
167. *Id.* at 328.
168. *Id.* at 332 (citations omitted).
169. 169 F.3d 330 (6th Cir. 1999).
170. *Id.* at 332.
171. 487 F.3d 323 (6th Cir. 2007).
172. *Id.* at 327–28, 343.

stention—would be inappropriate because the law was not subject to a narrowing construction that would bring it within the requirements of the Federal Constitution.[173] While the Sixth Circuit ultimately decided to neither certify nor abstain, the case is notable because it discusses certification, rather than abstention, when faced with *Pullman*-type issues.[174] Indeed, the court did not even cite to *Pullman*; it relied instead on *Arizonans for Official English* to establish the framework for analyzing whether to certify.[175]

In *Planned Parenthood Cincinnati Region v. Strickland*,[176] the Sixth Circuit heard a constitutional challenge to an Ohio law that prohibited the "off-label" use of the drug commonly known as RU-486.[177] After finding that *Pullman* considerations were at play (again, without actually citing to *Pullman*), the court found certification appropriate.[178] Notably, the opinion implied that the decisions whether to abstain or to certify required the same analysis[179] and that once a court found that either was appropriate, it should generally choose certification.[180] Indeed, although the court initially analyzed whether certification or abstention was appropriate, the bulk of the analysis dealt solely with certification.[181]

G. Seventh Circuit

The Seventh Circuit Court of Appeals has only decided a few cases involving *Pullman* considerations or certification since *Arizonans for Official English*. In *Shegog v. Board of Education of the City of Chicago*,[182] a case brought under federal question jurisdiction that involved primarily state law claims, the court considered certifying the state law issues to the Illinois courts but ultimately opted for an abstention-like remedy.[183] The court remanded the case to the district court with instruc-

---

173. *Id.* at 343.
174. *Id.*
175. *Id.*
176. 531 F.3d 406 (6th Cir. 2008).
177. *Id.* at 408–09.
178. *Id.* at 410–11. The court relied primarily upon the Supreme Court's decisions in *Bellotti v. Baird*, 428 U.S. 132 (1976), and *Arizonans for Official English*. *Id.*
179. *See id.*:

> Where a statutory interpretation is at issue, the United States Supreme Court has instructed the federal courts to employ *certification or abstention* if the "unconstrued state statute is susceptible of a construction by the state judiciary which might avoid in whole or in part the necessity for federal constitutional adjudication, or at least materially change the nature of the problem."

(emphasis added) (quoting *Bellotti*, 428 U.S. at 146–47).
180. *See id.*
181. *See id.* at 410–12. For another case in which the Sixth Circuit has opted for certification without consideration of the *Pullman* factors, or, indeed, citation to *Pullman*, see *American Booksellers Foundation for Free Expression v. Strickland*, 560 F.3d 443, 444–45, 447 (6th Cir. 2009) (certifying the construction of state law restricting use of the Internet for distribution of sexually explicit materials to juveniles).
182. 194 F.3d 836 (7th Cir. 1999).
183. *Id.* at 840.

tions to determine whether the federal claim was sound.[184] If the claim was sound, the district court was instructed to "provide appropriate relief and then either relinquish supplemental jurisdiction . . . or defer further proceedings until the state courts have had an opportunity to address the main state-law question."[185] If the claim was not sound, the district court was instructed to dismiss the case and direct the plaintiffs to seek redress in state court.[186] The court preferred not to certify the question because it was concerned that "[c]ertification may present the question in a needlessly abstract way . . . or short-circuit the normal process of decision by the state's intermediate appellate courts."[187] By definition, certification involves a question presented in a somewhat abstract manner to the state's highest court (thus, "short-circuiting" the trial-level and intermediate state courts), but this was the precise procedure endorsed in *Arizonans for Official English*. The court's deviation from the *Arizonans* framework was likely related to its appraisal that the lone federal claim was weak. However, by not citing this as an explanation for its decision not to certify, the Seventh Circuit's decision stands at odds with the *Arizonans* framework.[188]

H. Eighth Circuit

The Eighth Circuit Court of Appeals has also heard few cases involving either *Pullman* or *Arizonans for Official English* since *Arizonans* was decided. The case law is sparse: only one Eighth Circuit case contains any substantive discussion of abstention or certification.

In that case, *List v. County of Carroll*,[189] the Eighth Circuit heard a challenge to the district court's decision to dismiss claims alleging Missouri state practices violated the Americans with Disabilities Act and the Rehabilitation Act.[190] The Eighth Circuit affirmed the dismissal, but on different grounds than the district court's decision.[191] Although the court noted that at least some of the claims were appropriate for *Pullman* abstention,[192] its analysis was brief and did not mention the possibility of certification.

I. Ninth Circuit

In the years since *Arizonans for Official English*, the Ninth Circuit Court of Appeals' *Pullman*-type cases have been inconsistent and inadequately explained. Many of the court's abstention or certification deci-

184. *Id.*
185. *Id.* (citation omitted).
186. *Id.*
187. *Id.*
188. *See supra* note 81.
189. 240 F. App'x 155 (8th Cir. 2007) (per curiam).
190. *Id.* at 156.
191. *Id.*
192. *Id.* at 157.

sions go unpublished—a pattern that does not advance the development of robust and resilient case law or provide guidance to lower courts. More than any other circuit, therefore, the Ninth Circuit's approach to the issue has been inconsistent and poorly reasoned.

In the 1998 case *San Remo Hotel v. City & County of San Francisco*,[193] the Ninth Circuit considered a takings challenge to a San Francisco ordinance regulating hotel rooms.[194] The court abstained after concluding that the takings question could be avoided by allowing the state courts to interpret the ordinance at issue.[195] The court did not, however, mention certification as an option. This may have been because the California certification rule states that a federal district court may not certify a question of state law to the California Supreme Court; that privilege is limited to the U.S. Supreme Court, federal courts of appeal, or any other state's court of last resort.[196] However, due to the Ninth Circuit's silence on this issue in *The San Remo Hotel*—as well as its other *Pullman*-type cases involving questions of California law—it is impossible to know whether this rule impacted the court's decision. Nor is it possible to discern how a litigant might induce the Ninth Circuit to certify a question on appeal that had prompted a district court to abstain based on an unclear question of California law.

In an unpublished opinion in 2000, *Cucamongans United for Reasonable Expansion v. City of Rancho Cucamonga*,[197] the Ninth Circuit considered an appeal of the district court's decision to abstain pursuant to *Pullman*.[198] The appellants requested that the circuit court certify the state law questions.[199] The circuit court addressed this request in two short sentences: "The decision whether to certify a question of state law to a state supreme court rests in the sound discretion of the federal court. . . . After reviewing the briefs, record, and relevant law in this case, we conclude that certification to the California Supreme Court is not appropriate."[200] The court of appeals held that the lower court had not abused its discretion in abstaining. That same year, in two other unpublished cases, the Ninth Circuit reviewed lower court decisions to abstain—both of which it ultimately upheld—but did not mention certification as an option in either case.[201]

193. 145 F.3d 1095 (9th Cir. 1998).
194. *Id.* at 1098.
195. *Id.* at 1104.
196. CAL. R. CT. 8.548.
197. 210 F.3d 382, 2000 WL 61312 (9th Cir. 2000) (unpublished table decision).
198. *Id.* at *1.
199. *Id.*
200. *Id.* at *2 (citation omitted).
201. Resist the List v. Selecky, 242 F.3d 383, 2000 WL 1507524, at *1 (9th Cir. 2000) (unpublished table decision); Santa Clara County Corr. Peace Officers' Ass'n, Inc. v. Bd. of Supervisors, 225 F.3d 663, 2000 WL 734387, at *2 (9th Cir. 2000) (unpublished table decision).

In a lengthy opinion examining the legality of a city ordinance regulating rental housing, the Ninth Circuit concluded in *Columbia Basin Apartment Association v. City of Pasco*[202] that *Pullman* abstention was appropriate as applied to some of the plaintiffs, and that *Younger* abstention was appropriate for the remaining plaintiffs.[203] Although the plaintiffs in *Columbia Basin* suggested certification as an alternative to abstention, the majority opinion did not address their request.[204] However, the dissent disagreed with the majority's finding that *Younger* and *Pullman* abstention applied to the parties in this case and suggested that a "better course of action would be . . . to certify the doubtful questions of state law to the Washington Supreme Court."[205]

In an unpublished opinion from 2002, *Fetish & Fantasy Halloween Ball, Inc. v. Ahern Rentals, Inc.*,[206] the Ninth Circuit found that *Pullman* abstention was appropriate for a question involving the constitutionality of a Nevada prejudgment attachment statute.[207] While the court did not discuss whether certification would have been available in the case, it did note that abstention was particularly appropriate because there was already a "pending state-court action between [the parties that] provide[d] an opportunity for the Nevada courts to construe the attachment statute."[208]

In another opinion from 2002, *Parents Involved in Community Schools v. Seattle School District, No. 1*,[209] a Ninth Circuit panel decided to certify a question regarding whether a school district's use of race to assign students to different schools violated Washington's anti-discrimination statute.[210] Although the court did not reference the *Pullman* decision, the case clearly involved *Pullman* considerations inasmuch as the Ninth Circuit was motivated to certify the question so as to "avoid making federal constitutional decisions unless and until necessary."[211] The case had been before the Ninth Circuit previously,[212] and while the analysis as to whether to certify was brief (and contained no mention of *Pullman* or abstention), its brevity may have been due to the panel's familiarity with the details and extensive history of the case.

---

202. 268 F.3d 791 (9th Cir. 2001).
203. *Id.* at 794.
204. *See id.* at 805.
205. *Id.* at 810 (Tashima, J., dissenting).
206. 45 F. App'x 585 (9th Cir. 2002).
207. *Id.* at 588.
208. *Id.* at 587.
209. 294 F.3d 1085 (9th Cir. 2002).
210. *Id.* at 1086–87. This case has a lengthy procedural history. The Supreme Court ultimately reversed the Ninth Circuit's substantive constitutional holdings in 2007. *See* Parents Involved in Cmty. Sch. v. Seattle Sch. Dist., No. 1, 551 U.S. 701, 710–11 (2007).
211. *Parents Involved in Cmty. Sch.*, 294 F.3d at 1092 (quoting Clark v. City of Lakewood, 259 F.3d 996, 1016 n.12 (9th Cir. 2001)).
212. *Id.* at 1086.

In an unpublished 2004 decision, *Lueck v. Nevada Judicial Ethics & Election Practices Commission*,[213] a Ninth Circuit panel found that the plaintiff likely did not have standing and remanded the case "with instructions to dismiss [the plaintiff's] complaint without prejudice to any right he may have to seek review by the Nevada Supreme Court."[214] The court reasoned that whether the plaintiffs had standing turned on a question of state law, a determination the court considered a proper basis for *Pullman* abstention.[215] Nowhere in the opinion, however, did the court mention the option of certification.

In *Smelt v. County of Orange*,[216] a challenge to California's law prohibiting same-sex marriage, the Ninth Circuit held that the district court had not abused its discretion when it abstained.[217] At issue was whether the California law violated the California Constitution and the U.S. Constitution.[218] The court did not address certification as an option. However, in its analysis, the court emphasized that abstention was particularly appropriate in this case because very similar cases were pending in the state court system and were likely to reach the California Supreme Court.[219]

Most recently, the Ninth Circuit heard a challenge to a city's practice of denying health insurance to former employees.[220] In *Doyle v. City of Medford*,[221] a group of retirees alleged that the city's practice violated Oregon law and the Fourteenth Amendment's Due Process Clause.[222] The plaintiffs' due process claim was based on the contention that Oregon law conferred on them a property interest in health insurance benefits during their retirement.[223] The Ninth Circuit noted an ambiguity in the Oregon statute at issue that had not been interpreted or resolved by the Oregon courts.[224] The Ninth Circuit, in deciding to certify, wrote simply: "without guidance from the Oregon Supreme Court about what the [ambiguous] phrase . . . means in this context or . . . [what] the legislature intended, we are unable to decide the federal constitutional question accurately."[225] Interestingly, the court then cited Oregon case law governing certification of questions from federal courts to the Oregon Supreme Court and noted that one factor bearing on the certification decision is whether the issues in the case "could implicate the doctrine of

213. 106 F. App'x 552 (9th Cir. 2004).
214. *Id.* at 555.
215. *Id.* at 554.
216. 447 F.3d 673 (9th Cir. 2006).
217. *Id.* at 686.
218. *Id.* at 676–77.
219. *Id.* at 681.
220. Doyle v. City of Medford, 565 F.3d 536, 537 (9th Cir. 2009).
221. *Id.*
222. *Id.* at 537.
223. *Id.* at 538.
224. *Id.* at 541.
225. *Id.* at 542.

'*Pullman* abstention.'"[226] The court devoted considerable space to its discussion of whether *Pullman* factors were at play in this case—ultimately concluding that they were—and relied on both *Arizonans for Official English* and an Oregon case for the proposition that certification would avoid many of the inefficiencies associated with abstention.[227] This analysis, while thorough, was misplaced. To analyze the appropriateness of abstention and certification, along with the relative merits of each, only after having decided to certify, puts the cart before the horse. The court presents its analysis of abstention in an attempt to persuade the Oregon Supreme Court to accept and answer the certified question, without using any of this analysis to support its initial decision to seek state court guidance.[228]

J. Tenth Circuit

While the Tenth Circuit Court of Appeals has heard relatively few cases involving *Pullman* considerations since *Arizonans for Official English*, it has provided one well-reasoned and comprehensive opinion that discusses when to certify and when to abstain.

Kansas Judicial Review v. Stout[229] involved a First Amendment challenge to the Kansas Code of Judicial Conduct canon that prohibits judicial candidates from making certain types of campaign pledges.[230] The Tenth Circuit chose to certify the "important and unsettled questions of state law" underlying the claim to the Kansas Supreme Court.[231] In deciding to certify, the court first looked to whether the case was an appropriate candidate for *Pullman* abstention.[232] Without coming to any firm conclusions as to whether the case justified abstention, the court noted the problems inherent in abstention—especially the effect of abstention-related delay in First Amendment cases—and the Supreme Court's expression of a preference for certification, rather than abstention.[233] It endorsed certification as "consistent with our duties to avoid passing on the constitutionality of a statute where possible," especially when it is a state statute that has yet to be interpreted by the state's highest court.[234] The court concluded that the decision to certify rests within a federal court's discretion, but that "where statutory interpretation is at issue, the touchstone of our certification inquiry is whether the state sta-

226. *Id.* at 543 (citing W. Helicopter Servs., Inc. v. Rogerson Aircraft Corp., 811 P.2d 627, 630 (Or. 1991)).

227. *Id.* at 543–44 (citing Arizonans for Official English v. Arizona, 520 U.S. 43, 75 (1997); W. Helicopter Servs., Inc., 811 P.2d at 632).

228. *Id.* at 543 ("We recognize that the court takes into account several discretionary factors when deciding whether to accept a question for certification.").

229. 519 F.3d 1107 (10th Cir. 2008).

230. *Id.* at 1111.

231. *Id.*

232. *Id.* at 1118–19.

233. *Id.* at 1119.

234. *Id.*

tute is readily susceptible of an interpretation that 'would avoid or substantially modify the federal constitutional challenge to the statute.'"[235]

K. Eleventh Circuit

Like the Tenth Circuit, the Eleventh Circuit Court of Appeals has heard few cases involving *Pullman* considerations since *Arizonans for Official English* was decided. The circuit's sole opinion dealing with the question of when to certify rather than abstain, however, provides an accessible and detailed framework to guide lower courts.

In *Pittman v. Cole*,[236] the Eleventh Circuit considered a First Amendment challenge to the "enforcement of advisory opinions promulgated by the Alabama Judicial Inquiry Commission and the Alabama State Bar's Office of General Counsel" that would have prevented judicial candidates from answering election-related questionnaires.[237] The district court had abstained under the *Pullman* doctrine.[238] The Eleventh Circuit concluded that the decision to abstain was erroneous, even as it noted that "the district court was correct in recognizing that unsettled issues of state law could well shape if not moot the plaintiffs' federal constitutional claims."[239] Though the case involved *Pullman* considerations, the Eleventh Circuit held that the district court abused its discretion because it failed to consider several prudential factors, including "'delay, cost, doubt as to the adequacy of state procedures[,] . . . the existence of factual disputes, and [whether] the case has already been in litigation for a long time,'" as well as the type of claim at issue.[240] The court emphasized that *Pullman* abstention would rarely be appropriate in cases involving First Amendment challenges because delay is particularly injurious to First Amendment rights.[241] The Eleventh Circuit then remanded the case to the district court with instructions to certify the questions of state law to the Alabama Supreme Court.[242] In so doing, it emphasized that where abstention might have proven "problematic," certification simplified the analysis and allowed the federal court to request clarification from state courts.[243]

IV. RECOMMENDATIONS: TWO APPROACHES TO PROPER CERTIFICATION

The inconsistency within and across the circuits has created a confusing backdrop for litigants and for lower courts considering whether to certify a question of state law in cases involving a federal constitutional

235. *Id.* at 1119–20 (quoting Bellotti v. Baird, 428 U.S. 132, 148 (1976)).
236. 267 F.3d 1269 (11th Cir. 2001).
237. *Id.* at 1273, 1276.
238. *Id.* at 1285.
239. *Id.*
240. *Id.* at 1286–87 (quoting Duke v. James, 713 F.2d 1506, 1510 (11th Cir. 1983)).
241. *Id.* at 1290.
242. *Id.* at 1291.
243. *Id.* at 1290.

challenge to a state law. The Supreme Court's decision in *Arizonans for Official English* made clear that lower courts must, at a very minimum, consider certification when they otherwise would have abstained under *Pullman*. *Arizonans* did not, however, provide an analytical framework to guide lower courts as they determine whether to abstain, to certify, or to decide for themselves the state law issues. As the previous section shows, certain circuits—most notably the Second Circuit, the Tenth Circuit, and the Eleventh Circuit—have developed a clear, coherent framework for determining whether to abstain or certify. Other circuits, such as the First Circuit and the Ninth Circuit, have failed to provide a consistent, thorough, and transparent approach to this issue. This inconsistency is problematic for numerous reasons. Here, however, I simply note that dramatic inconsistencies can lead to unfair results and, just as importantly, a public perception of arbitrary decision-making.[244]

In this section, I suggest two approaches to deciding whether to abstain or certify in *Pullman*-type cases. Both approaches are consistent with the Supreme Court's cases on the issue of certification in cases involving federal constitutional challenges to state laws. The first approach, derived from the Second Circuit's cases, combines the *Pullman* doctrine with recent certification decisions from the Supreme Court to create a list of factors that each court should consider as it determines whether certification or abstention is appropriate. The second approach is a two-step analysis that modifies the Second Circuit framework by breaking the inquiry into two separate prongs.

A. The Tunick *Analysis*

The first approach involves the framework devised by Judge Calabresi in *Tunick v. Safir*.[245] In *Tunick*, Judge Calabresi emphasized that,

---

244. *See* Stephen B. Burbank, *Judicial Independence, Judicial Accountability, and Interbranch Relations*, 95 GEO. L.J. 909, 915 (2007):

> Research suggests that diffuse support [of the courts] is linked to legitimizing messages about the courts, such as those that highlight the role of precedent and the rule-of-law ideal, and that it is adversely affected by delegitimizing messages, such as those that frame court decisions simply in terms of results (for example, the message that *Bush v. Gore* decided the 2000 election).

Sarah E. Ricks, *The Perils of Unpublished Non-Precedential Federal Appellate Opinions: A Case Study of the Substantive Due Process State-Created Danger Doctrine in One Circuit*, 81 WASH. L. REV. 217, 217 (2006):

> Doctrinal divergence between the Third Circuit's binding and non-precedential opinions has undermined the predictive value of precedential state-created danger decisions, creating an obstacle to settlement at both the trial and appellate levels. In turn, district courts' unpredictable application of the non-precedential opinions has undermined the critical appellate functions of ensuring that like cases are treated alike, that judicial decisions are not arbitrary, and that legal issues resolved at the appellate level need not be relitigated before the district courts.

See also Robert S. Thompson, *Legitimate and Illegitimate Decisional Inconsistency: A Comment on Brilmayer's* Wobble, or the Death of Error, 59 S. CAL. L. REV. 423, 432–35 (1986) (describing normative objections to judicial inconsistency).

245. 209 F.3d 67 (2d Cir. 2000).

though *Arizonans for Official English* should be read as instructing the use of certification "in more instances than had previously been thought appropriate," *Arizonans* does not hold that a court "must certify whenever (a) a plaintiff raises a federal constitutional challenge to state law in federal court, (b) the state's highest court has not interpreted the statute, and (c) the constitutional question could conceivably be avoided by some saving interpretation."[246] His opinion concluded that any discussion of the appropriateness of certification in federal constitutional challenges to state statutes must rely on *Pullman* "for guidance,"[247] and that, "[g]iven the shared goals of *Pullman* abstention and of the device of certification, the factors counseling the former are also suggestive of when the latter is desirable."[248] *Arizonans for Official English* and the Second Circuit's right-to-die cases, Judge Calabresi reasoned, add a "gloss" to the *Pullman* factors, but do not undermine or displace the *Pullman* analysis.[249]

From these various strands of case law, Judge Calabresi produces six factors that a court should consider when determining whether to certify a case involving federal constitutional challenges to state law: (1) whether the state statute has already been interpreted by the state's highest court; (2) how important the issue is and whether it is likely to recur; (3) whether serious constitutional issues may be avoided by certain interpretations of the state statute; (4) the extent to which certification would "render federal constitutional decisions unnecessary or . . . ensure that they are inescapably before the federal court"; (5) whether the state law issue implicates core issues of state governance and how a potential federal decision would impact federalism concerns; and (6) how the certification-related delay would impact the claimed federal rights.[250]

These six factors will clarify whether certification is the appropriate action or whether the federal court should instead construe the state statute without certifying a question to the state court. The *Tunick* analysis convincingly adapts the Supreme Court's cases into an accessible six-factor test that, if uniformly applied by federal appellate courts, would impose much-needed consistency in this area of law. What these six factors do not explicitly address are those situations in which certification may be appropriate, but, for reasons addressed in Part II of this article, abstention may be even more appropriate. In other words, the *Tunick* analysis does not account for instances in which the federal court should seek state court guidance, but where certification is unavailable or inappropriate because of doctrinal reasons (under *Pennhurst* other forms of

246. *Id.* at 73–74.

247. *Id.* at 74 ("[A]lthough *Pullman* abstention involves problems that certification may avoid or reduce, it still remains the doctrine whose purpose is most proximate to that of certification in cases concerning the federal constitutional validity of state laws.").

248. *Id.* at 75.

249. *Id.*

250. *Id.* at 81.

abstention) or pragmatic reasons (such as the existence of a similar case pending before the state supreme court).

B. The Two-Step Analysis

The two-step approach would require a court considering certification or abstention to engage in a straightforward two-step analysis: first, whether the case involves *Pullman* considerations, and second, whether any factors exist that weigh against the presumption of certification. Under this approach, a court would first determine whether or not a case involves *Pullman* considerations.[251] During this first step, the court would temporarily ignore both the possibility of certification and the question of whether abstention would be the best option. Instead, the court would focus solely on the existence of *Pullman* considerations. A court would analyze whether there is: (1) a difficult constitutional issue involving core issues of state governance that (2) could be avoided by a narrowing construction of a (3) state law that has yet to be construed by the state's highest court.[252] From this analysis, the court will have assessed whether seeking guidance—either certification or abstention—from the state courts would be appropriate.

After ascertaining whether the *Pullman* factors are satisfied, the court would proceed to determine whether any factors exist that weigh against the presumption of certification and suggest that the court either abstain or decide the issue on its own. If any of the prudential or doctrinal considerations discussed in Part II of this article were present, then the court should balance them with the advantages inherent in certification.[253] For example, a court might decide to abstain, rather than certify, because the case involves *Pennhurst* considerations that would render the

251. While I argue here that a court determining whether to certify on *Pullman* grounds should first determine whether the threshold requirements for *Pullman* abstention are met, I do not suggest that certification is only appropriate if these threshold requirements are met. Indeed, certification is appropriate in many situations in which abstention would most definitely not be appropriate. *See* FALLON ET AL., *supra* note 42, at 1201. This article deals only with the limited set of cases in which a court is considering *Pullman* abstention and has the option to certify rather than abstain. In non-*Pullman* certification, a court determining whether to certify should not use the abstention analysis as its starting point. *See* Deborah J. Challener, *Distinguishing Certification from Abstention in Diversity Cases: Postponement Versus Abdication of the Duty to Exercise Jurisdiction*, 38 RUTGERS L.J. 847, 885 (2007) (arguing that certain limitations on abstention are not applicable to certification, and that courts need not apply principles of abstention when determining whether to certify).

252. *See, e.g.*, Nicholson v. Scoppetta, 344 F.3d 154, 168 (2d Cir. 2003) ("We need not certify or abstain unless 'the [state] statute is fairly subject to an interpretation which will render unnecessary or substantially modify the federal constitutional question' . . . . Nor are we obliged to avoid constitutional questions that are not 'serious.'" (alteration in original) (citations omitted)).

253. *See* Pittman v. Cole, 267 F.3d 1269, 1289 (11th Cir. 2001) (emphasizing certification's benefits and finding that certification was appropriate where abstention was not).

court unable to grant the requested relief,[254] or because parallel proceedings in the state courts are already under way.[255]

The mere presence of one of the factors outlined in Part II should not, however, end the inquiry. Whether the existence of a certain factor should lead the court to abstain, rather than certify, will differ from case to case. For instance, the presence of parallel state court proceedings (assuming the inapplicability of *Younger*) might not justify abstention if the case involved a First Amendment challenge on the grounds that the delay that accompanies abstention can cause particular harm to First Amendment freedoms.[256] This result might not be the same in a case that does not involve the First Amendment. In the end, this analysis is essentially a balancing test: the court must weigh the various pragmatic and doctrinal considerations and determine, on balance, whether certification or abstention is more appropriate. Of course, the Supreme Court's statements favoring certification must inform a court's balancing act; in the end, certification should have a thumb on the scale.

This two-step analysis is not altogether unfamiliar. Indeed, some courts already employ similar analyses in their approach to cases involving *Pullman* considerations.[257] For instance, in *Pittman v. Cole*,[258] the Eleventh Circuit first determined that, while *Pullman* factors were present in the case, factors such as the likely impact of abstention-related delay and the interim relief granted to plaintiffs made abstention inappropriate.[259] The court then considered certification, noting that "[a]lthough we conclude that the district court erred by abstaining in the manner that it did in this case, we agree with it that there exist important unsettled issues of state law that are likely to shape, alter, or moot the federal constitutional issues raised by the plaintiffs' claims."[260] The court concluded that certification was appropriate in light of the advantages of certification, as well as the Supreme Court's endorsement of the procedure in *Arizonans for Official English*.[261]

---

254. *See, e.g.*, Univ. of Utah v. Shurtleff, 252 F. Supp. 2d 1264, 1285 (D. Utah 2003) (abstaining under the *Pullman* doctrine, in part, because the *Pennhurst* doctrine precluded the court from addressing the state law claims in the case).

255. *See, e.g.*, Rivera-Feliciano v. Acevedo-Vilá, 438 F.3d 50, 61–62 (1st Cir. 2006) (finding *Pullman* abstention appropriate because the same issues were pending before the Supreme Court of Puerto Rico as a result of parallel state court proceedings).

256. *See, e.g.*, Kan. Judicial Review v. Stout, 519 F.3d 1107, 1119 ("Courts have been particularly reluctant to abstain in cases involving facial challenges on First Amendment grounds . . . in part because the delay caused by declining to adjudicate the issues could prolong the chilling effect on speech." (citation omitted)).

257. And, of course, other courts may be employing this balancing approach without making it explicit. But if they do not make their reasoning explicit, it is difficult for the case law to develop and for lower courts to know how they ought to approach the issue.

258. 267 F.3d 1269 (11th Cir. 2001).

259. *Id.* at 1285–88.

260. *Id.* at 1288.

261. *Id.* at 1288–91.

The two-step test takes after the *Tunick* analysis inasmuch as it incorporates the six *Tunick* factors. Both approaches would make clear whether certification is an appropriate course of action, and both tests would also bring a consistent analytic approach to this field of law. By separating the inquiry into two steps, however, the two-step framework provides more information. As with *Tunick*, the two-step approach helps the court determine whether certification is appropriate. But it also helps the court determine whether certain factors counsel in favor of abstention. It is important to not combine these steps, because the inquiries are fundamentally different: whether to ask for participation from the state courts versus how to facilitate state court participation.

The two-step structured analysis has the potential to bring clarity and even elegance to an area of law that has long been unwieldy and unnecessarily confusing. Imposing a degree of formality to the analysis would provide predictability and guidance to lower courts and parties.[262] By following the Supreme Court's ruling in *Arizonans for Official English* more closely, federal appellate courts would also protect the federalism values embodied in both *Arizonans* and *Pullman*. And, perhaps most relevant to practitioners and litigants, they will ensure that parties are spared the unnecessary delay and expense involved in abstention that the *Arizonans for Official English* decision sought to eliminate.

CONCLUSION

Certification and abstention play an important role in the complex and changing relationship between federal and state courts. With certification's rising frequency, abstention under the *Pullman* doctrine has necessarily waned. As this article has shown, however, the theory and reasoning underlying *Pullman* abstention remain essential to any analysis of certification in cases where previously a court would have considered abstaining under the *Pullman* doctrine. If the circuit courts are to develop a coherent and thorough approach to certification in cases where they previously would have abstained, they must look to *Pullman* to deter-

---

262. As discussed *supra* note 244, inconsistent or unexplained decision-making may result in arbitrary decisions (or a perception of arbitrariness) and may undermine the legitimacy of the court system. It also makes the job of federal district courts more difficult. *See also* Martha J. Dragich, *Will the Federal Courts of Appeals Perish if They Publish? Or Does the Declining Use of Opinions to Explain and Justify Judicial Decisions Pose a Greater Threat?*, 44 AM. U. L. REV. 757, 800 (1995) ("By failing to provide sufficient guidance to lower court judges, summary dispositions, selective publication practices, and noncitation rules undermine certainty, predictability, and fidelity to authority.").

Of course, even with a more formalistic framework, the federal courts retain discretion in determining whether to certify or abstain. *See* R.R. Comm'n of Tex. v. Pullman Co., 312 U.S. 496, 501 (1941) ("These cases reflect a doctrine of abstention appropriate to our federal system whereby the federal courts, 'exercising a wise discretion,' restrain their authority because of 'scrupulous regard for the rightful independence of the state governments' and for the smooth working of the federal judiciary." (quoting Cavanaugh v. Looney, 248 U.S. 453, 457 (1919); Di Giovanni v. Camden Fire Ins. Ass'n, 296 U.S. 64, 73 (1935)).

mine whether to involve the state courts and to *Arizonans for Official English* to determine how to seek state court guidance.

Courts considering a *Pullman*-type case should rely upon the *Tunick* framework or adopt the two-step analysis described in Part IV.B. The latter option is preferable, inasmuch as it addresses not just whether to seek state court guidance, but also what procedure to use to obtain state guidance. Nevertheless, both options will ensure compliance with the Supreme Court's decisions in *Pullman* and *Arizonans for Official English*. Following the guidance of these two cases will allow courts to develop a robust body of case law, provide a foundation for clear and thorough analysis, and lead to consistent results that reflect the federalism values at the heart of both decisions.

ARIZONA V. GANT: RETHINKING THE EVIDENCE-GATHERING JUSTIFICATION FOR THE SEARCH INCIDENT TO ARREST EXCEPTION, AND TESTING A NEW APPROACH

INTRODUCTION

The United States Supreme Court's decision in *New York v. Belton*[1] has been subject to recent criticism from scholars[2] and Supreme Court Justices[3] alike, calling for the Court to revisit its broad construction of the search incident to arrest exception to the Fourth Amendment's prohibition against warrantless searches and seizures in the vehicle-search context.

The Court seized the opportunity to reexamine *Belton* in *Arizona v. Gant*.[4] The *Gant* holding, while narrowing the scope of the search incident to arrest exception in some situations, extended its scope in others, perhaps straying from the tenets of the Fourth Amendment.[5]

Part I of this Comment recounts the inconsistent history of the search incident to arrest exception, and the state of the law prior to the *Gant* decision. Part II summarizes the facts, procedural history, and opinions of *Gant*. Part III analyzes *Gant*'s two-part holding, suggests an alternative rule for the second part, and tests the compatibility of the rule with another area of Fourth Amendment jurisprudence—the law relating to inventory searches. The Comment concludes that the Supreme Court should rethink its treatment of the Fourth Amendment's warrant requirement in the vehicle context.

I. BACKGROUND

A. Origin of the Search Incident to Arrest Exception

The Fourth Amendment protects individuals from unreasonable searches and seizures.[6] The search incident to arrest exception to the Fourth Amendment's prohibition against warrantless searches provides that, under certain circumstances, law enforcement officers in the field may conduct searches incident to a lawful arrest without a search war-

1. 453 U.S. 454 (1981).

2. *See, e.g.*, David S. Rudstein, Belton *Redux: Reevaluating* Belton*'s Per Se Rule Governing the Search of an Automobile Incident to an Arrest*, 40 WAKE FOREST L. REV. 1287, 1288 (2005).

3. *See, e.g.*, Thornton v. United States, 541 U.S. 615, 632 (2004) (Scalia, J., concurring).

4. 129 S. Ct. 1710 (2009).

5. *See* James J. Tomkovicz, *Divining and Designing the Future of the Search Incident to Arrest Doctrine: Avoiding Instability, Irrationality, and Infidelity*, 2007 U. ILL. L. REV. 1417, 1473–74 (2007).

6. U.S. CONST. amend. IV.

rant.[7] The exception originated in the dictum of *Weeks v. United States*.[8] Distinguishing *Weeks* from previous cases, the Court stated that the issue was "not an assertion of the right on the part of the government always recognized under English and American law, to search the person of the accused when legally arrested, to discover and seize the fruits or evidences of crime."[9]

Since *Weeks*, the Supreme Court has wrestled with a satisfactory definition of the scope of authority to search granted by the exception.[10] In that time the proper scope has repeatedly expanded and contracted.

B. United States v. Rabinowitz[11]

In *Rabinowitz*, federal agents, incident to a lawful arrest for forgery of stamps, searched a desk, safe, and file cabinet in the one-room office where the arrest was made.[12] The Court held that the search was reasonable under the search incident to arrest exception because the search was not "general or exploratory," but sought evidence relevant to the crime of arrest, which the agents had reason to believe was hidden in the office.[13] *Rabinowitz* allowed searches conducted under this evidence-gathering justification to extend beyond the person of the arrestee, and beyond the area immediately surrounding him.[14]

Justice Frankfurter's dissent warned that the search incident to arrest exception was meant to be narrow and based on necessity.[15] Further, and particularly relevant to the *Gant* decision, Justice Frankfurter noted that extending searches beyond the arrestee's person for the purpose of gathering evidence allowed searches without probable cause.[16] The dissent warned that this broad grant of authority to field officers, absent judicial scrutiny through the warrant-issuing process, could lead to a slippery slope and eventually result in a police state.[17]

---

7. *Gant*, 129 S. Ct. at 1716.
8. 232 U.S. 383, 392 (1914).
9. *Id.*
10. Tomkovicz, *supra* note 5, at 1421.
11. 339 U.S. 56 (1950).
12. *Id.* at 57–59.
13. *Id.* at 62–63.
14. *See id.* at 61.
15. *Id.* at 72 (Frankfurter, J., dissenting) (citing Robertson v. Baldwin, 165 U.S. 275, 281 (1897)).
16. *See id.* at 80:

> The purpose of the Fourth Amendment was to assure that the existence of probable cause as the legal basis for making a search was to be determined by a judicial officer *before arrest and not after*, subject only to what is necessarily to be excepted from such requirement. The exceptions cannot be enthroned into the rule.

(emphasis added).
17. *Id.* at 82 ("The progress is too easy from police action unscrutinized by judicial authorization to the police state.").

C. Chimel v. California[18]

In *Chimel*, police officers arrested Ted Chimel in his home, and proceeded to search the entire three-bedroom house without a search warrant.[19] The Supreme Court held the search was unreasonable, and crafted a narrower search incident to arrest rule that expressly overruled *Rabinowitz*.[20] *Chimel* held that "it is reasonable for the arresting officer to search the person" of the arrestee without a search warrant, as well as the "area 'within his immediate control.'"[21]

The Court recognized only two justifications supporting use of the search incident to arrest exception: (1) ensuring officer safety by preventing the arrestee from accessing weapons, and (2) thwarting the destruction of evidence by preventing access to such evidence.[22]

D. New York v. Belton[23]

In *Belton*, the Supreme Court applied the exception to arrests of vehicle occupants, and established a broad scope of police authority for vehicle searches.[24] The Court held that "when a policeman has made a lawful custodial arrest of the occupant of an automobile, he may, as a contemporaneous incident of that arrest, search the passenger compartment of that automobile."[25] The rule also applied to recent occupants[26] and allowed searches of containers within the passenger compartment, but excluded the trunk from the searchable area.[27]

The Court reasoned that, because no workable definition existed in the vehicle context for the "area within the immediate control of the arrestee,"[28] law enforcement officers needed a straightforward rule for efficient application of the search incident to arrest exception.[29]

In his dissenting opinion, Justice Brennan argued that the Court sacrificed Fourth Amendment principles to obtain a bright-line rule.[30] Justice Brennan attacked the majority's acceptance of a legal fiction that merely paid lip service to *Chimel*'s twin justifications of protecting offic-

---

18. 395 U.S. 752 (1969).
19. *Id.* at 754.
20. *Id.* at 768.
21. *Id.* at 763 (defining "within his immediate control" as "the area from within which he might gain possession of a weapon or destructible evidence").
22. *Id.*
23. 453 U.S. 454 (1981).
24. *See id.* at 460–61.
25. *Id.* at 460.
26. *See id.*
27. *Id.* at 461 & n.4 (defining "container" as "any object capable of holding another object").
28. *Id.* at 460.
29. *Id.* at 459–60 ("When a person cannot know how a court will apply a settled principle to a recurring factual situation, that person cannot know the scope of his constitutional protection, nor can a policeman know the scope of his authority.").
30. *Id.* at 468 (Brennan, J., dissenting) ("[T]he Court does a great disservice . . . to the policies underlying the Fourth Amendment . . . .").

ers and preserving evidence.[31] Justice Brennan noted that when officers arrest an occupant of a vehicle, "the justifications underlying *Chimel*'s limited exception to the warrant requirement cease to apply," because the passenger compartment cannot be in the "area within his immediate control" when the arrestee is physically restrained.[32]

Therefore, although the Court insisted its holding "in no way alters the fundamental principles established in the *Chimel* case,"[33] Justice Brennan argued the fiction that "the interior of a car is *always* within the immediate control of an arrestee who has recently been in the car" cannot be reconciled with *Chimel*.[34] Justice Brennan also questioned how much direction the supposed bright-line rule would actually supply, and contended that *Chimel*, unmodified, provided adequate guidance for law enforcement in the vehicle context.[35] Twenty-three years passed before the Court revisited the *Belton* decision.

E. Thornton v. United States[36]

Thornton affirmed and extended *Belton*, holding that an officer had the right to search an arrestee's vehicle incident to the arrest even when the officer's first contact with the arrestee occurred outside of the vehicle.[37] Because the arrestee was a "recent occupant" of the vehicle, the Court found the search reasonable.[38] However, five justices expressed dissatisfaction with the state of the law as governed by *Belton* in the search incident to arrest exception for motor vehicles.[39] In addition to the two dissenting justices, Justice O'Connor voiced her disapproval in a concurring opinion,[40] while Justice Scalia, joined by Justice Ginsburg, concurred only in the judgment.[41]

Justice Scalia's opinion attacked the fiction of using *Chimel*'s twin justifications to support the reasonableness of *Belton* searches.[42] Justice Scalia observed the risk that a handcuffed arrestee, secured in the back of a squad car, might obtain a weapon or destructible evidence from within his vehicle "was remote in the extreme."[43]

---

31. *See id.* at 466.
32. *Id.* at 466.
33. *Id.* at 460 n.3 (majority opinion).
34. *See id.* at 466 (Brennan, J., dissenting).
35. *Id.* at 471–72.
36. 541 U.S. 615 (2004).
37. *See id.* at 623–24 (holding that the arrestee was a "recent occupant" under *Belton* where the arrestee pulled into a parking lot and exited his vehicle before the arresting officer was able to pull him over, and the officer pulled in behind him, exited his vehicle, and initiated contact).
38. *Id.* ("So long as an arrestee is the sort of 'recent occupant' of a vehicle such as [the arrestee] was here, officers may search that vehicle incident to the arrest.").
39. *Id.* at 624, 632.
40. *Id.* at 624 (O'Connor, J., concurring).
41. *Id.* (Scalia, J., concurring).
42. *See id.* at 631.
43. *Id.* at 625.

The concurrence advocated applying the evidence-gathering justification from the overruled *Rabinowitz* case in the vehicle context.[44] Justice Scalia contended warrantless vehicle searches incident to arrest should be valid only "where it is reasonable to believe evidence relevant to the crime of arrest might be found in the vehicle."[45] The approach conspicuously required less than probable cause to perform a warrantless search incident to arrest.[46] Yet, the opinion gave several reasons why warrantless vehicle searches should require less than a showing of probable cause, including a reduced expectation of privacy in motor vehicles, and "heightened law enforcement needs" because the mobility of vehicles creates a greater risk that evidence within the vehicle will be lost.[47]

Thornton showed the readiness of five justices to alter *Belton*'s broad construction of the search incident to arrest exception as applied to vehicles. Five years after *Thornton*, the Supreme Court changed the scope of the exception in *Arizona v. Gant*.

II. *ARIZONA V. GANT*[48]

A. Facts

On August 25, 1999, Tucson police officers responded to a tip about an illegal drug enterprise operating in a residence.[49] Rodney Gant answered the door, identified himself, and told the officers that the owner of the house would return later.[50] A records check on Gant showed an outstanding warrant for his arrest for driving with a suspended license.[51]

The officers returned the same evening, and had finished arresting and securing a man and a woman near the house when Gant pulled into the residence's driveway.[52] Gant parked, stepped out of his vehicle, and an officer arrested and handcuffed Gant between ten and twelve feet from the vehicle.[53] After placing the handcuffed Gant in the backseat of a squad car, the officers searched his vehicle and found a gun and a bag of cocaine in a jacket on the backseat.[54]

B. Procedural History

Facing charges of possession of a narcotic drug for sale and possession of drug paraphernalia, Gant moved to suppress the evidence found

44. *See id.* at 629 (citing United States v. Rabinowitz, 339 U.S. 56, 60–64 (1950)).
45. *Id.* at 632.
46. *See id.*
47. *Id.* at 630–32 (citing Wyoming v. Houghton, 526 U.S. 295, 303–04 (1999)).
48. 129 S. Ct. 1710 (2009).
49. *Id.* at 1714.
50. *Id.* at 1714–15.
51. *Id.* at 1715.
52. *Id.*
53. *Id.*
54. *Id.*

during the search of his car, contending the search violated his Fourth Amendment right to be free from unreasonable searches.[55] The trial court found the search reasonable and denied Gant's motion.[56]

The Arizona Supreme Court reversed, holding the search unreasonable under the Fourth Amendment because neither of *Chimel*'s twin justifications for a search incident to arrest existed once Gant was handcuffed in the back of a squad car.[57] The Arizona Supreme Court stated it did not purport to "reconsider" *Belton*, and explained that because the rationales of officer safety and evidence preservation were not present to justify the search, the search incident to arrest violated the existing Fourth Amendment law.[58] The Court distinguished *Thornton*, noting that the defendant in *Thornton* never raised the argument that the underlying justifications for the exception were absent, and only contended that it did not apply because he was outside of his vehicle when he first encountered the police.[59] The State of Arizona petitioned the United States Supreme Court for certiorari, and the Court, responding to persistent calls to revisit the *Belton* decision,[60] granted the petition.[61]

C. Majority Opinion

Justice Stevens delivered the opinion of the Court, which held that the officers conducted an unreasonable search of Gant's vehicle.[62] The Court held that *Chimel*'s rationale[63] "authorizes police to search a vehicle incident to a recent occupant's arrest only when the arrestee is unsecured and within reaching distance of the passenger compartment at the time of the search."[64] The Court also held that police may search a vehicle incident to a recent occupant's arrest where it is "'reasonable to believe evidence relevant to the crime of arrest might be found in the vehicle.'"[65]

Because the *Gant* rule treats the exception differently depending on the crime of arrest,[66] the rule is best explained by splitting the holding into its two parts. Professor James Tomkovicz articulates the classifications particularly well, separating crimes of arrest into "evidentiary" and

---

55. *Id.*
56. *See id.*
57. *Id.* at 1715–16.
58. State v. Gant, 162 P.3d 640, 643 (Ariz. 2007).
59. *Id.* at 644.
60. *See, e.g.*, Rudstein, *supra* note 2, at 1288 ("Given the dissatisfaction with the *Belton* rule expressed by five Justices in *Thornton*, it is an appropriate time to evaluate again the holding and reasoning of *Belton*.").
61. Arizona v. Gant, 128 S. Ct. 1443 (2008) (mem.).
62. *Gant*, 129 S. Ct. at 1724.
63. Chimel v. California, 395 U.S. 752, 763 (1969) (recognizing that ensuring officer safety and preventing destruction of evidence are the two justifications underlying use of the search incident to arrest exception).
64. *Gant*, 129 S. Ct. at 1719.
65. *Id.* (quoting Thornton v. United States, 541 U.S. 615, 632 (2004) (Scalia, J., concurring)).
66. *Id.*

"nonevidentiary" offenses.[67] An arrest for a nonevidentiary offense supplies no reason to believe that evidence specific to the offense of arrest might be found in the vehicle, while an arrest for an evidentiary offense *does* create a reasonable belief that such evidence might be found.[68]

1. Arrests for Nonevidentiary Offenses

The Court distinguished *Belton* and *Thornton* by classifying driving with a suspended license as the type of offense "for which police could not expect to find evidence in the passenger compartment of Gant's car."[69] If the offense of arrest is nonevidentiary, only *Chimel*'s twin justifications of ensuring officer safety and preserving evidence validate warrantless searches of vehicles incident to arrest of an occupant or recent occupant, and only then if "the arrestee is unsecured and within reaching distance of the passenger compartment at the time of the search."[70] The Court's holding effectively overruled *Belton* for nonevidentiary offenses, establishing a narrow application of the exception.[71]

The Court rejected the State's argument that *Belton* searches are reasonable even where there is no possibility that the arrestee might access the passenger compartment, and explained that the new rule "correctly balances law enforcement interests, including the interest in a bright-line rule, with an arrestee's limited privacy interest in his vehicle."[72] The Court reasoned that a broad reading of *Belton* gives police too much authority to conduct warrantless searches when no reasonable basis exists to believe evidence of the offense might be found in the vehicle.[73] The Court characterized police authority granted under *Belton* to rummage through a person's private effects as "the central concern underlying the Fourth Amendment."[74]

2. Arrests for Evidentiary Offenses

As proposed by Justice Scalia in *Thornton*, the Court adopted the evidence-gathering justification for vehicle searches incident to arrests

---

67. Tomkovicz, *supra* note 5, at 1418.

68. *Gant*, 129 S. Ct. at 1719:

> In many cases, as when a recent occupant is arrested for a traffic violation, there will be no reasonable basis to believe the vehicle contains relevant evidence. But in others, including *Belton* and *Thornton*[, both drug offenses], the offense of arrest will supply a basis for searching the passenger compartment of an arrestee's vehicle and any containers therein.

(citations omitted).

69. *Id.*

70. *Id.*

71. *See id.* ("To read *Belton* as authorizing a vehicle search incident to every recent occupant's arrest would thus untether the rule from the justifications underlying the *Chimel* exception . . . . [W]e reject this reading of *Belton*.").

72. *Id.* at 1720.

73. *Id.*

74. *Id.*

for evidentiary offenses.[75] Justice Stevens suggested drug offenses are evidentiary offenses, referencing the crimes at issue in *Belton* and *Thornton*.[76] Both cases concerned drug offenses, which Justice Stevens contended created a reasonable belief that evidence relevant to those crimes could have been found in the arrestees' respective vehicles.[77] The Court also held an arrest for an evidentiary crime supplies a "basis for searching the passenger compartment and any containers therein."[78] By so holding, the Court retained the physical scope of the *Belton* search,[79] but only in situations giving rise to an evidence-gathering justification to conduct the search.[80]

D. Concurring Opinion

Justice Scalia wrote separately in *Gant*, despite the Court having adopted his proposed rule from *Thornton*. Justice Scalia lamented the "charade" of retaining *Chimel*'s justifications in the car search context.[81] As an alternative, he advocated adopting *only* an evidence-gathering justification for warrantless vehicle searches incident to arrests of recent occupants.[82] Justice Scalia reasoned that allowing searches under *Chimel*'s rationale where arrestees are unsecured invites officers to leave the scene unsecured.[83] Interestingly, in *Thornton*, decided just five years earlier, he quickly dismissed the same argument.[84]

E. Dissenting Opinion

Justice Alito's dissent, in which Chief Justice Roberts and Justice Kennedy joined, and Justice Breyer joined in part, attacked the Court's insufficient support for its departure from *stare decisis*.[85] The dissent also questioned adopting Justice Scalia's proposed rule from his concurring

75. *See id.* at 1719; Thornton v. United States, 541 U.S. 615, 632 (2004) (Scalia, J., concurring).

76. *See Gant*, 129 S. Ct. at 1719.

77. *Id.*

78. *Id.*

79. *Compare id.* (holding searches of a vehicle's passenger compartment reasonable when incident to evidentiary arrests), *with* New York v. Belton, 453 U.S. 454, 460–61 & n.4 (1981) (defining the area of the "passenger compartment" as encompassing the compartment itself as well as all containers within the compartment, including glove compartments, consoles, luggage, boxes, bags, and clothing, but not the trunk).

80. *Gant*, 129 S. Ct. at 1719.

81. *Id.* at 1724–25 (Scalia, J., concurring).

82. *Id.* at 1725.

83. *Id.* at 1724–25 ("[T]his standard . . . leaves much room for manipulation, inviting officers to leave the scene unsecured . . . in order to conduct a vehicle search.").

84. Thornton v. United States, 541 U.S. 615, 627 (2004) (Scalia, J., concurring) ("[I]f an officer leaves a suspect unrestrained nearby just to manufacture authority to search, one could argue that that search is unreasonable *precisely because* the dangerous conditions justifying it existed only by virtue of the officer's failure to follow sensible procedures.").

85. *See Gant*, 129 S. Ct. at 1727 (Alito, J., dissenting).

opinion in *Thornton* without independent explanation of its origin or rationale.[86]

The dissent noted law enforcement's considerable reliance on *Belton*, as evidenced by the fact that it was taught in police academies.[87] The dissent argued that the *Belton* rule offers a more workable alternative than *Gant*'s rule.[88] Justice Alito contended that *Gant*, by narrowing the application of *Chimel*'s justifications,[89] creates a "'perverse incentive for an arresting officer to prolong the period during which the arrestee is kept in an area where he could pose a danger to an officer.'"[90]

Finally, Justice Alito questioned the part of the new rule that requires "reason to believe" rather than probable cause.[91] The dissent concluded that it would simply apply *Belton*, and reverse the holding of the Arizona Supreme Court.[92]

III. ANALYSIS

A. Gant's *Abandonment of* Belton *for Nonevidentiary Offenses*

The first part of the *Gant* holding, narrowing the scope of vehicle searches incident to arrests for nonevidentiary offenses, presents two concerns: (1) that the new rule abandoned a workable bright-line rule crafted in *Belton*, thwarting the efficiency of law enforcement;[93] and (2) that the change in the rule's wording undermines the justifications it purports to advance.

1. *Gant* Correctly Abandoned *Belton*'s Rule for Nonevidentiary Offenses

a. *Belton* Failed to Provide an Adequate Bright-Line Rule

In *Gant*, Justice Stevens noted the inconsistency of lower courts in applying *Belton* to recent occupants of vehicles.[94] The disparities resulted from *Belton*'s failure to provide any instruction in determining the required spatial relationship between arrestee and vehicle at the time of

86. *Id.* at 1726.

87. *Id.* at 1728.

88. *Id.* at 1729 ("[S]erious problems will also result from the second part of the Court's new rule, which requires officers making roadside arrests to determine whether there is reason to believe that the vehicle contains evidence of the crime of arrest.").

89. *See id.* at 1719 (majority opinion).

90. *Id.* at 1730 (Alito, J., dissenting) (quoting United States v. Abdul-Saboor, 85 F.3d 664, 669 (D.C. Cir. 1996)). *But see* Thornton v. United States, 541 U.S. 615, 627 (2004) (Scalia, J., concurring) ("The weakness of this argument is that it assumes that, one way or another, the search must take place. But conducting a *Chimel* search is not the Government's right; it is an exception—justified by necessity—to a rule that would otherwise render the search unlawful.").

91. *Gant*, 129 S. Ct. at 1731 (Alito, J., dissenting) ("Why . . . is the standard for this type of evidence-gathering search 'reason to believe' rather than probable cause?").

92. *Id.* at 1732.

93. *See id.* at 1729.

94. *Id.* at 1721 (majority opinion).

arrest, and the temporal relationship between the time of arrest and time of search.[95] In other words, how close to the vehicle did an arrest have to occur to authorize a search of the vehicle incident to the arrest, and how soon after the arrest must police search the vehicle?

In 1986, the Supreme Court of Connecticut held a search unreasonable that continued after police removed the arrestee from the scene.[96] Eight years later, the First Circuit Court of Appeals upheld a search continuing after police removed the arrestee from the scene.[97] Other decisions have wrestled with finding a required spatial relationship between arrestee and vehicle at the time of arrest.[98] A rule creating opposite results in similar situations offers no workable bright line.

In *Gant*, the Court specified the required spatial[99] and temporal[100] relationships for a reasonable vehicle search incident to arrest for a non-evidentiary offense. The majority correctly found that *Belton's* bright line only worked to confuse lower courts and breed inconsistent decisions.[101]

b. Balancing the Interest in Efficient Law Enforcement Against the Protections of the Fourth Amendment

In *Thornton*, Justice O'Connor characterized lower courts' application of *Belton* as treating "the ability to search a vehicle incident to the arrest of a recent occupant as a police *entitlement*."[102] In *Gant*, the Court stated that the fact that law enforcement relied on *Belton*'s broad construction of search authority to the point of treating it as an entitlement could not outweigh the constitutionally guaranteed individual privacy interest.[103]

---

95. *See id.* at 1720–21 & nn.6–7; *see also Thornton*, 541 U.S. at 622 (establishing that the exception applies when an officer makes initial contact with the arrestee outside of the vehicle).

96. State v. Badgett, 512 A.2d 160, 169 (Conn. 1986).

97. United States v. Doward, 41 F.3d 789, 793 & n.3 (1st Cir. 1994).

98. *Compare* United States v. Caseres, 533 F.3d 1064, 1072 (9th Cir. 2008) (declining to apply *Belton* when police approached arrestee after he had left his vehicle and reached his residence), *with* Black v. State, 810 N.E.2d 713, 713–14, 716 (Ind. 2004) (applying *Belton* when arrestee was apprehended inside a shop while his vehicle was parked outside), *and* Rainey v. Commonwealth, 197 S.W.3d 89, 94–95 (Ky. 2006) (upholding a vehicle search incident to the arrest of its recent occupant, where police apprehended the occupant fifty feet from the vehicle).

99. *Gant*, 129 S. Ct. at 1719 (holding the *Chimel* rationale applies only when an arrestee is "within reaching distance of the passenger compartment").

100. *Id.* (holding the arrestee must be "within reaching distance of the passenger compartment *at the time of the search*" (emphasis added)).

101. *See id.* at 1720–21 & nn.6–7. *But see id.* at 1729 (Alito, J., dissenting) ("The first part of the new rule . . . reintroduces the same sort of case-by-case, fact-specific decisionmaking that the *Belton* rule was adopted to avoid.").

102. Thornton v. United States, 541 U.S. 615, 624 (2004) (O'Connor, J., concurring) (emphasis added).

103. *Gant*, 129 S. Ct. at 1723 (majority opinion).

Justice Alito based his dissent on the fact that police officers had relied on *Belton*'s rule for over twenty-five years.[104] However, length of reliance on a rule should not matter if that rule allows activity which unlawfully diminishes a constitutional protection.[105] In *Gant*, the Court properly concluded that the interest in efficient law enforcement cannot outweigh the importance of the Fourth Amendment's protection against unreasonable searches and seizures for nonevidentiary crimes.

2. The Change from "Area into Which an Arrestee Might Reach"[106] to "Within Reaching Distance"[107] Undermines *Chimel*'s Twin Justifications

Under *Gant*, the arrestee must be "unsecured and *within reaching distance* of the passenger compartment at the time of the search" in order for the officer to search the vehicle based on *Chimel*'s justifications.[108] In a situation where an officer is outnumbered by unsecured arrestees outside the reaching distance of the vehicle,[109] the officer can reasonably fear the arrestees may attempt to overpower the officer and access weapons or destructible evidence within the vehicle. But a court applying the *Gant* rule might hold a vehicle search under these facts unreasonable, because the arrestees are not in "reaching distance of the passenger compartment."[110] *Chimel*'s phrasing[111] offers more room to argue that the passenger compartment *is* in the area within the immediate control of the arrestees in such a situation.

In a footnote in *Gant*, the Court recognizes the unlikelihood of a situation where an officer cannot complete an arrest by securing the arrestee or arrestees in handcuffs,[112] but argues that a search incident to arrest would be reasonable in these situations.[113] Yet, the Court's rephrasing of *Chimel*'s rule, and its qualifying footnote, unnecessarily muddle a once-clear rule statement. Even considering the footnote's recognition of situa-

---

104. *Id.* at 1728 (Alito, J., dissenting).

105. *See id.* at 1723 (majority opinion):

> If it is clear that a practice is unlawful, individuals' interest in its discontinuance clearly outweighs any law enforcement 'entitlement' to its persistence . . . . [N]one of the dissenters in *Chimel* or the cases that preceded it argued that law enforcement reliance interests outweighed the interest in protecting individual constitutional rights so as to warrant fidelity to an unjustifiable rule.

106. Chimel v. California, 395 U.S. 752, 763 (1969).

107. *Gant*, 129 S. Ct. at 1719.

108. *Id.* (emphasis added).

109. This was the case in *Belton*: there, the arresting officer, lacking enough handcuffs to secure the four arrestees, instructed the arrestees to stand in four different areas of the road while he conducted a search of their persons and then the passenger compartment. *See* New York v. Belton, 453 U.S. 454, 456 (1981).

110. *Gant*, 129 S. Ct. at 1719.

111. *Chimel*, 395 U.S. at 763 (the area "within [the arrestee's] immediate control").

112. *Gant*, 129 S. Ct. at 1719 n.4 (citing 3 WAYNE R. LAFAVE, SEARCH & SEIZURE: A TREATISE ON THE FOURTH AMENDMENT § 7.1 (4th ed. 2004)).

113. *See id.* (holding that a search is reasonable where an officer is unable to perform an arrest because a real possibility exists that the arrestee might gain access to the passenger compartment).

tions analogous to *Belton*, a strong argument exists that "reaching distance" simply means an arm's length.

The *Gant* rule creates potential situations where an arresting officer cannot search a vehicle, even where the facts implicate the justifications for the exception. The *Chimel* construction of the rule offers a workable solution to unusual fact patterns like *Belton*, while recognizing any exception to the Fourth Amendment's warrant requirement must be narrow and well-delineated. Returning to *Chimel*'s original wording best serves the exception's justifications because it governs the situation, however unlikely, where an arrestee outside reaching distance of the vehicle threatens either officer safety or preservation of evidence.[114]

B. A Suggested Approach for Applying the Exception to Evidentiary Crimes

The second part of *Gant*'s holding allows police to perform a warrantless search of a vehicle incident to an arrest when a reasonable belief exists that evidence relevant to the offense of arrest could be found in the vehicle.[115] As the dissent noted, the new rule for evidentiary crimes substitutes "reason to believe" for probable cause as the standard for performing a search under the Fourth Amendment.[116]

While valid reasons support establishing a broad scope for evidence-gathering searches incident to arrest in the vehicle context,[117] an expansive construction of the exception raises equally valid concerns.[118] The relevant consideration is whether the government interest in effective law enforcement outweighs the individual privacy interest in being free from unreasonable searches under the Fourth Amendment.

Perhaps less intrusive methods exist to further the government's interests. If the Court considers the interest of ensuring successful prosecutions important enough to disregard constitutionally protected rights in

---

114. *See Belton*, 453 U.S. at 471 (Brennan, J., dissenting) ("[R]elevant factors [in determining the area within the arrestee's immediate control] would surely include the relative number of police officers and arrestees . . . and the ability of the arrestee to gain access to a particular area or container.").

115. *Gant*, 129 S. Ct. at 1719 (citing Thornton v. United States, 541 U.S. 615, 632 (2004) (Scalia, J., concurring)).

116. *Id.* at 1731 (Alito, J., dissenting).

117. *See Thornton*, 541 U.S. at 631 (Scalia, J., concurring) (advancing a "reduced expectation of privacy" and "heightened law enforcement needs" as justifications for allowing evidence-gathering searches); *see also Gant*, 129 S. Ct. at 1720 ("[A] motorist's privacy interest in his vehicle is less substantial than in his home . . . ."); Wyoming v. Houghton, 526 U.S. 295, 304 (1999) ("[T]he 'ready mobility' of an automobile creates a risk that the evidence or contraband will be permanently lost while a warrant is obtained." (quoting California v. Carney, 471 U.S. 386, 390 (1985))).

118. *See* United States v. Rabinowitz, 339 U.S. 56, 70 (1950) (Frankfurter, J., dissenting) (discussing the importance of adhering to the Fourth Amendment's purpose that any invasion into one's privacy must be judicially sanctioned); Tomkovicz, *supra* note 5, at 1471–72 (characterizing the allowance of warrantless vehicle searches incident to arrest under an evidence-gathering justification without a probable cause requirement as a worrisome grant of unchecked power to law enforcement officers in the field).

pursuit of evidence, it should at least consider a rule that serves these interests *without* violating the reasonableness mandate of the Fourth Amendment. This mandate is served either by procuring a warrant based on sufficient probable cause, or acting within the scope of an exception sufficiently tied to the justifications upon which it is based.[119]

1. The Suggested Rule for Applying the Vehicle Search Incident to Arrest Exception to Evidentiary Crimes

The *Gant* rule works on an assumption that, because of the mobility of vehicles, there is no time to reasonably procure a search warrant for the vehicle an officer desires to search. An ideal rule would preserve the Fourth Amendment's search warrant requirement, and still recognize an evidence-gathering justification in the vehicle context. The rule would read: Before searching a vehicle incident to the arrest of an occupant or recent occupant for the purpose of gathering evidence relevant to the crime of arrest, an arresting officer "must secure and use search warrants wherever reasonably practicable."[120]

The rule would recognize that, because of the ready mobility of vehicles, there will often be instances where obtaining a search warrant is not reasonably practicable. Where obtaining a search warrant is not reasonably practicable—and there is reason to believe that relevant evidence inside the vehicle will be transported or destroyed—an officer may temporarily restrict access to the vehicle[121] until a search warrant can be obtained.

This rule attempts to protect the government's interest in gathering evidence to ensure successful prosecutions, and to satisfy the Fourth Amendment's purpose that an objective, judicial mind determines whether the situation justifies an intrusion into the arrestee's privacy.[122] Existing Fourth Amendment law regarding impounded vehicles, however, necessitates further analysis to determine the plausibility of the rule in situations where obtaining a search warrant would *not* be reasonably practicable.

---

119. *See* Katz v. United States, 389 U.S. 347, 357 (1967) ("[S]earches conducted outside the judicial process, without prior approval by judge or magistrate, are *per se* unreasonable under the Fourth Amendment—subject only to a few specifically established and well-delineated exceptions."). *But see* Groh v. Ramirez, 540 U.S. 551, 572–74 (2004) (Thomas, J., dissenting) (arguing that the Fourth Amendment does not contain an explicit warrant requirement, and requires only "reasonableness" in the search).

120. Trupiano v. United States, 334 U.S. 699, 705 (1948) (citing Johnson v. United States, 333 U.S. 10, 14, 15 (1948); Taylor v. United States, 286 U.S. 1, 6 (1932); Go-Bart Importing Co v. United States, 282 U.S. 344, 358 (1931); Carroll v. United States, 267 U.S. 132, 156 (1925)).

121. Impounding the vehicle would constitute a temporary restriction of access.

122. *See McDonald*, 335 U.S. at 455 ("[The warrant requirement was created] so that an objective mind might weigh the need to invade . . . privacy in order to enforce the law.").

2. Impounded Vehicles and the Fourth Amendment: The Inventory Search

a. Background of Inventory Searches of Impounded Vehicles

The authority of police to inventory lawfully impounded vehicles is a well-defined exception to the Fourth Amendment's search warrant and probable cause requirements.[123] However, inventory searches must be reasonable under the circumstances, complying with the Fourth Amendment's prohibition against unreasonable searches.[124] A vehicle must be properly impounded, giving police lawful custody over it, for an inventory search to be reasonable.[125]

The Supreme Court recognizes three administrative, "community caretaking functions"[126] as the justifications for inventory searches of impounded vehicles: (1) to protect the owner's property within the vehicle; (2) to protect police from claims or disputes regarding lost or stolen property; and (3) to protect the police and the public from potential danger.[127] The presence of any of these justifications authorizes a warrantless inventory search of an impounded vehicle.[128]

Officers must perform an inventory search in good faith and in furtherance of the above-mentioned administrative functions, as the inventory "must not be a ruse for a general rummaging in order to discover incriminating evidence."[129] Conducting the search in accordance with standardized police procedure normally satisfies the good faith requirement.[130] A court will likely find an inventory search reasonable so long as standardized procedures allow the scope of the search employed by an officer.[131] And, even if an officer suspects the existence of evidence of a

123. *See, e.g.*, South Dakota v. Opperman, 428 U.S. 364, 367–76 (1976) (identifying the "inherent mobility of automobiles" and a diminished privacy interest in vehicles—as opposed to the home—as the two reasons for the exception to the warrant requirement); *see also* Colorado v. Bertine, 479 U.S. 367, 371 (1987); Illinois v. Lafayette, 462 U.S. 640, 643 (1983).

124. *See Bertine*, 479 U.S. at 373–74.

125. *See Opperman*, 428 U.S. at 371 n.5; *see also* Woodford v. State, 752 N.E.2d 1278, 1281 (Ind. 2001) (holding impoundments are proper and recognizing lawful custody where the impoundment is part of the "administrative caretaking functions of the police," or where "authorized by state statute").

126. *Opperman*, 428 U.S. at 368 (quoting Cady v. Dombrowski, 413 U.S. 433, 441 (1973)).

127. *Id.* at 368–69.

128. *See Bertine*, 479 U.S. at 373 (holding that where police took a vehicle to a secure, lighted storage facility, the fact that there existed little risk of theft or vandalism of property in the vehicle did not eliminate the need for an inventory because the other two justifications for an inventory search were still present).

129. Florida v. Wells, 495 U.S. 1, 4 (1990).

130. *Bertine*, 479 U.S. at 374–75.

131. *See Wells*, 495 U.S. at 4 ("[P]olicies of opening all containers or of opening no containers [within the vehicle] are unquestionably permissible . . . .").

crime in the vehicle, it does not preclude the police from inventorying the vehicle so long as they follow the standardized procedure.[132]

The scope of a reasonable vehicle inventory includes a search of any area of the vehicle that serves the government community caretaking interests.[133] Supreme Court decisions hold the scope of a reasonable inventory search extends to an unlocked glove compartment,[134] and to containers in the passenger compartment.[135] Therefore, the scope of an inventory search at least matches the scope of a search incident to arrest, and perhaps exceeds it,[136] assuming *Gant* retained *Belton*'s prohibition on trunk searches incident to arrest.[137]

b. Plausibility of the Suggested Approach in Light of the Inventory Search Doctrine

Assuming, for the purpose of testing the suggested rule, that an impoundment incident to the arrest of a vehicle occupant constitutes a proper impoundment and authorizes an inventory of the vehicle, the suggested rule likely fails. The existence and pervasiveness of the inventory search exception to the Fourth Amendment's warrant requirement hinders the suggested rule's effectiveness where "obtaining a search warrant is not reasonably practicable."[138] The inventory search exception frustrates the purpose of the suggested rule because Supreme Court decisions recognize the inventory search as a routine police activity incident to impoundment,[139] and if a vehicle in fact contains evidence of a crime, police will inevitably discover it via a warrantless search regardless of the reasons behind that search. However, while the existence of the inventory search may negate any exclusionary remedy for evidence found during an unconstitutional application of the search incident to arrest exception, it does not preclude victims of such unconstitutional searches

132. *See* United States v. Petty, 367 F.3d 1009, 1013 (8th Cir. 2004) ("[P]olice 'may keep their eyes open for potentially incriminating items that they may discover in the course of an inventory search, as long as their sole purpose is not to investigate a crime . . . .'" (quoting United States v. Marshall, 986 F.2d 1171, 1176 (8th Cir. 1993))); *see also* State v. Ture, 632 N.W.2d 621, 629 (Minn. 2001) (holding that, if an officer has an investigatory motive for conducting an inventory search, the search remains reasonable so long as it is coupled with a non-investigatory motive).

133. *See* South Dakota v. Opperman, 428 U.S. 364, 376 n.10 (1976).

134. *Id.* (holding the search of a glove compartment reasonable in order to prevent theft, and to protect the police and public if "vandals" found a firearm).

135. *See Wells*, 495 U.S. at 4 (holding that a standardized police policy allowing officers to open all containers in the course of an inventory search would be reasonable).

136. *See, e.g.*, United States v. Duncan, 763 F.2d 220, 223 (6th Cir. 1985) (holding that an inventory search extending to a trunk was reasonable because of the interest in police protecting themselves from false claims of stolen or lost property). *But see* United States v. Wilson, 636 F.2d 1161, 1165 (8th Cir. 1980) ("[T]he needs of the government in conducting an inventory search may be ordinarily accomplished without the serious intrusion into the locked trunk of an automobile.").

137. New York v. Belton, 453 U.S. 454, 461 n.4 (1981).

138. *See supra* Part III.B.1.

139. *See Opperman*, 428 U.S. at 392 (Marshall, J., dissenting) ("The court's result authorizes indeed it appears to require the routine search of nearly every car impounded.").

from bringing suit for violations of their constitutional rights.[140] Therefore, while the suggested rule may protect the Fourth Amendment right of individuals to be free from unreasonable searches in this setting, it would likely have little effect on any exclusionary remedy available to a criminal defendant.

i. An Inventory Search Will Reveal Any Evidence that Officers Would Have Found by Conducting a Vehicle Search Incident to Arrest

Because the inventory search is virtually routine procedure, and because the scope of the search matches that of the vehicle search incident to arrest exception, the result is the same as if *Gant*'s rule for evidentiary crimes remained unmodified: police will search the vehicle without a warrant and discover any existing evidence which would have been discovered by a search incident to arrest.

The suggested rule only precludes warrantless vehicle searches incident to arrest under an evidence-gathering justification. However, once impounded, police will still search the same vehicle without a warrant under administrative justifications.

The suggested rule would only pay lip service to the Fourth Amendment's objectives.[141] It would prohibit warrantless searches under an evidence-gathering justification, but would not change the fact that the same warrantless search, with the same result, will take place under administrative justifications. In reality, is a defendant awaiting prosecution for a drug offense concerned whether police found the drugs in his car under an evidentiary or administrative rationale? Or is the defendant simply concerned that police found the drugs in his car?

ii. The Inevitable Discovery Rule

The inevitable discovery rule works as a catch-all in an exclusionary analysis, tying together the exceptions to the warrant requirement. The prosecution invokes the inevitable discovery rule as a challenge to exclusion of evidence unconstitutionally obtained by the government.[142] If the prosecution can establish by a preponderance of the evidence that the evidence would have inevitably been discovered by lawful means, then the evidence is admissible.[143]

The inevitable discovery doctrine defeats any exclusionary goal of the suggested rule. Picture a situation requiring impoundment because

140. 42 U.S.C. § 1983 (1996).

141. McDonald v. United States, 335 U.S. 451, 455 (1948) (defining the objective of the Fourth Amendment as interposing "an objective mind" between the citizenry and law enforcement to determine whether the interest in enforcing the law justifies an invasion into privacy).

142. *See* Nix v. Williams, 467 U.S. 431, 442–43 (1984).

143. *See id.* at 442–44.

obtaining a warrant is not reasonably practicable. Impoundment will eventually result in an inventory of the vehicle. Any evidence officers find will be admissible under the inevitable discovery rule because an inventory would have eventually revealed the evidence lawfully in the absence of the illegal search. While the illegal search would potentially open officers to civil liability,[144] the damage to the vehicle owner is done if contraband is actually found.

The suggested rule[145] serves the Fourth Amendment's purpose of allowing a judicial mind to determine the validity of police invasions into privacy in terms of establishing a defined constitutional protection under the search incident to arrest exception. However, one way or another, it allows the same result it was designed to curtail—a police invasion of privacy supported neither by a search warrant, nor by probable cause, in the form of an inventory search. Similarly, under the *Gant* rule, it appears that where an arrest of a vehicle occupant requires an impoundment, the prosecution will have access to evidence obtained by an officer exceeding the lawful scope of the search.

3. The Inventory Search Doctrine Is Insufficiently Tied to the Justifications for its Existence, and Therefore Must Be Reconsidered

The inventory search is an unnecessarily intrusive—and fairly ineffective—method of furthering the caretaking needs on which it is based, making it the type of thinly veiled general rummage that the Fourth Amendment sought to eliminate.[146] Although the Supreme Court holds that an inventory search does not require a warrant or probable cause,[147] the three administrative needs on which the inventory search is based fail to justify an exception to the warrant requirement.[148] Also, alternative methods exist to serve these needs and still adhere to the Fourth

144. For example, a claim brought pursuant to 42 U.S.C. § 1983 (2006).

145. *See supra* Part III.B.1:

> Before searching a vehicle incident to the arrest of an occupant or recent occupant for the purpose of gathering evidence relevant to the crime of arrest, an arresting officer 'must secure and use search warrants wherever reasonably practicable.' . . . Where obtaining a search warrant is not reasonably practicable, and there is reason to believe that a danger exists that relevant evidence inside the vehicle will be transported or destroyed, an officer may temporarily restrict access to [impound] the vehicle until a search warrant can be obtained.

(quoting Trupiano v. United States, 334 U.S. 699, 705 (1948)).

146. *See* New York v. Belton, 453 U.S. 454, 464 (1981) (Brennan, J., dissenting) ("[I]n determining whether to grant an exception to the warrant requirement, courts should carefully consider the facts and circumstances of each search and seizure, focusing on the reasons supporting the exception . . . ." (citing Sibron v. New York, 392 U.S. 40, 59 (1968))); *see also* Terry v. Ohio, 392 U.S. 1, 19 (1968) ("The scope of the search must be 'strictly tied to and justified by' the circumstances which rendered its initiation permissible." (quoting Warden v. Hayden 387 U.S. 294, 310 (1967) (Fortas, J., concurring))).

147. *See* South Dakota v. Opperman, 428 U.S. 364, 371 n.5 (1976).

148. *Id.* at 389 (Marshall, J., dissenting).

Amendment by protecting individuals from arbitrary invasions of privacy by law enforcement.[149]

a. Protection of the Vehicle Owner's Property

Protecting property within an impounded vehicle from theft and vandalism appears to be a benevolent reason for conducting an inventory search.[150] However, this rationale ignores owners who do not wish to have their vehicles searched for this purpose.[151] Further, without probable cause[152] of the presence of property in the vehicle worthy of protection, the search becomes a mere rummage through the property of another. This type of search does not further any caretaking objective and derives its only support from a general curiosity about the contents of the vehicle. This for-your-own-protection rationale[153] for the vehicle inventory search exception, in the absence of an owner's consent, should not outweigh the individual privacy interest implicated in the search.

As an alternative to inventory searches, perhaps police can increase security around impound lots in order to better protect the property within the vehicles. And, although *Opperman* contends that additional security measures—such as posting a guard around impound lots—could be "prohibitively expensive" for smaller jurisdictions,[154] this unproven prediction hardly qualifies as a circumstance of "such exigency that, as a practical necessity, rigorous enforcement of the warrant requirement is impossible."[155]

b. Protection of Police from False Claims of Stolen or Lost Property

South Dakota v. Opperman, the seminal case upholding warrantless inventories of impounded vehicles, advanced the protection of police from false claims of stolen or lost property as one justification for the exception.[156] But, as the dissent noted, South Dakota state law absolved the police in *Opperman* from any obligation other than inventorying items in plain view and locking the car, thus obviating any further police action for the purpose of protecting themselves from claims of stolen

149. *Contra* Illinois v. Lafayette, 462 U.S. 640, 647 (1983) ("The reasonableness of any particular governmental activity does not necessarily or invariably turn on the existence of alternative 'less intrusive' means.").

150. *See Opperman*, 428 U.S. at 369.

151. *Id.* at 392 (Marshall, J., dissenting) ("[I]t is obvious that not everyone whose car is impounded would want it to be searched.").

152. *See* Terry v. Ohio, 392 U.S. 1, 21 (1968) ("[I]n justifying the particular intrusion the police officer must be able to point to specific and articulable facts which, taken together with rational inferences from those facts, reasonably warrant that intrusion.").

153. *See Opperman*, 428 U.S. at 389 (Marshall, J., dissenting).

154. *Id.* at 379 (Powell, J., concurring).

155. *Id.* at 367 (majority opinion) (citing Carroll v. United States, 267 U.S. 132, 153–54 (1925)).

156. *Id.* at 369.

property.[157] Even if other jurisdictions lack similar provisions, the *Opperman* decision, upon which later Supreme Court holdings are based,[158] erred in reasoning that the search protected police from false claims of theft. Inventory searches do not effectively protect police from false claims[159]: a claimant can still assert that an officer stole an item before the inventory occurred, or that officers intentionally omitted an item's presence from their records.[160]

In the *Opperman* dissent, Justice Marshall offered a more effective, less intrusive method of security from false claims.[161] Justice Marshall suggested placing seals on the trunk and doors of the vehicle, where an unbroken seal signified that the car was unopened during police custody.[162] This too leaves police open to false claims, as owners might claim police broke a seal, and then replaced it. However, if the police used individually numbered seals, provided the numbers to the arrestee, and placed the seals on the vehicle in his presence, the arrestee would have no claim of stolen property if the seal remains unbroken.

c. Protection of Police and the Public from Potential Danger

While protection from potential danger undoubtedly represents a legitimate interest, the Supreme Court's broad construction of the exception allows inventory searches where no specific circumstances indicate the presence of a specific danger.[163] Specific facts should be required to implicate this justification because impoundment alone creates no reasonable belief of danger.[164] While situations implicating a safety rationale may arise, it cannot logically justify the search of *every* impounded vehicle.[165] However, the Supreme Court accepts this rationale, even where the "danger" is invented.[166] Using the *Opperman* Court's logic allowing searches based on the unsubstantiated presence of danger, what

157. *See* S.D. CODIFIED LAWS § 43-39-11 (2009) (implying that when police are in possession of an impounded car, they are acting as gratuitous depositors); State v. Opperman, 228 N.W.2d 152, 159 (S.D. 1975) (holding that as gratuitous depositors, police are protected from civil tort claims if they remove objects from the vehicle in plain view, close the windows, and lock the doors).

158. *See, e.g.*, Florida v. Wells, 495 U.S 1, 4 (1990); Colorado v. Bertine, 479 U.S. 367, 374–75 (1987); Illinois v. Lafayette, 462 U.S. 640, 647 (1983).

159. *Opperman*, 428 U.S. at 378–79 (Powell, J., concurring) ("It is not clear . . . that inventories are a completely effective means of discouraging false claims . . . .").

160. *Id.* at 379.

161. *See id.* at 391 n.10 (Marshall, J., dissenting).

162. *Id.*

163. *See, e.g.*, *Wells*, 495 U.S. at 2 (arrest for driving under the influence); *Bertine*, 479 U.S. at 367 (arrest for driving under the influence); *Opperman*, 428 U.S. at 365 (police impounded vehicle for parking violations).

164. *Opperman*, 428 U.S. at 378 (Powell, J., concurring) ("Except in rare cases, there is little danger associated with impounding unsearched automobiles.").

165. *Id.* at 389–90 (Marshall, J., dissenting) ("Even aside from the actual basis for the police practice in this case, however, I do not believe that any blanket safety argument could justify a program of routine searches of the scope permitted here.").

166. *See id.* at 376 n.10 (majority opinion) (holding a police inventory search of a vehicle impounded for a traffic violation reasonable for the purpose of protecting the public from vandals who might find a firearm).

is to stop police from inventorying parked cars on the street in the off chance that one of them contains a weapon?[167]

An argument exists that an arrest for an evidentiary crime and subsequent impoundment of the arrestee's vehicle provide probable cause that the vehicle contains evidence relevant to the crime of arrest. Combined with other specific facts offering a certain location and a certain item to search for, this is likely true. However, without specific facts showing that a specific item within a vehicle presents a danger, the exception becomes unhinged from the safety rationale, and reverts to a general rummage, unconnected to the reason for its existence.

4. A Suggested Course of Action

In reconsidering the inventory search, the Court should heed the advice of Justice Scalia: If a search is for an evidence-gathering purpose, then the Court should at least admit that purpose, and then test it against the Fourth Amendment.[168] But if the inventory search exception exists to serve the needs listed in *Opperman*, then the Court should narrowly tailor the grant of police authority to fulfill those needs. The above-mentioned alternatives[169] are viable ways to further the two interests in protecting vehicle owners from theft and protecting police from false claims of stolen property. The safety rationale simply cannot logically or constitutionally justify warrantless, broad-scope searches of *every* impounded vehicle.[170]

Adopting the suggested rule for evidentiary crimes and using less intrusive alternatives[171] to address administrative needs provides a workable framework. It acknowledges the importance of administrative interests—and the interests in gathering evidence to aid prosecution of suspected criminals—yet remains loyal to the purpose of the Fourth Amendment by adhering more closely to its warrant requirement. The increased costs these rules would place on police departments are concededly a burden. But this burden does not outweigh the importance of protecting the privacy of citizens from arbitrary rummages by government officials.

---

167. *See id.* at 391 n.8 (Marshall, J., dissenting) ("If this asserted rationale justifies search of all impounded automobiles, it must logically also justify the search of [a]ll automobiles, whether impounded or not, located in a similar area, for the argument is not based on the custodial role of the police.").

168. Thornton v. United States, 541 U.S. 615, 631 (2004) (Scalia, J., concurring) ("[I]f we are going to continue to allow *Belton* searches . . . we should at least be honest about why we are doing so.").

169. *See supra* Part III.B.3.a–b.

170. *Opperman*, 428 U.S. at 389–91 (Marshall, J., dissenting).

171. *See supra* Part III.B.3.a–b (suggesting increased security at impound lots and the use of numbered seals as alternatives that would protect property within vehicles and protect police from false claims).

CONCLUSION

The Supreme Court's decision in *Gant* cuts two ways. First, it returns to a narrow search incident to arrest exception for nonevidentiary crimes, although its wording might result in confused application of the rule by lower courts. Second, and more troubling, it reflects a continuing trend placing higher value on successful prosecution of suspected criminals than on the intent of the Fourth Amendment's framers. The *Gant* rule applied to evidentiary crimes in the vehicle context, by adopting an evidence-gathering justification with no probable cause requirement, effectively removes the judicial mind from the process and grants officers in the field the discretion to determine whether an invasion of privacy is justified. This is exactly what the framers sought to prevent.[172]

While the rule suggested above provides an alternative to warrantless vehicle searches incident to arrest, it does so only until tripping over the next exception to the warrant requirement. The existence and breadth of the inventory search exception virtually ensure that, under the proposed improvement, warrantless search is inevitable.

In the context of vehicle searches after the arrest of an occupant, the exceptions to the Fourth Amendment's warrant requirement have enveloped the rule. After an arrest, it is difficult to imagine a situation where the vehicle search would take place under authority of a search warrant. Whether the search is conducted incident to an arrest or as an inventory, the paths circumventing that set forth by the Fourth Amendment have become the most traveled. The Court's treatment of the Fourth Amendment in the vehicle context deserves a better explanation than the oft-repeated diminished privacy interest in vehicles,[173] because "[t]he word 'automobile' is not a talisman in whose presence the Fourth Amendment fades away."[174] Courts must avoid treating the Fourth Amendment as a "nuisance, [and] a serious impediment in the war against crime,"[175] and instead temper the need for effective law enforcement with the directives issued in the Bill of Rights.

To be sure, other freedoms in the Constitution might engender more support than Fourth Amendment protection from unreasonable search and seizure—the individuals who invoke the protection are often accused criminals. But, to quote Justice Frankfurter, "[I]t is precisely because the appeal to the Fourth Amendment is so often made by dubious characters that its infringements call for alert and strenuous resistance."[176]

---

172. Arizona v. Gant, 129 S. Ct. 1710, 1720 (2009) ("[T]he central concern underlying the Fourth Amendment [is] the concern about giving police officers unbridled discretion to rummage at will among a person's private effects.").

173. *See, e.g.*, United States v. Chadwick, 433 U.S. 1, 12–13 (1977).

174. Coolidge v. New Hampshire, 403 U.S. 443, 461 (1971).

175. Harris v. United States, 331 U.S. 145, 157 (1947) (Frankfurter, J., dissenting).

176. *Id.* at 156.

*Jason Hermele**

* J.D. Candidate, 2011. I thank the Law Review staff and editorial board for all their help, especially J.D. Schneider for his valuable suggestions and revisions. Also, thanks to Professor Justin Marceau for his thoughtful advice. Finally, thank you to my family and friends for your support.

WINTER V. NATIONAL RESOURCES DEFENSE COUNCIL: ENABLING THE MILITARY'S ONGOING ROLLBACK OF ENVIRONMENTAL LEGISLATION

INTRODUCTION

Congress passed the National Environmental Policy Act ("NEPA")[1] in an effort to encourage environmental responsibility in federal agency decision-making.[2] Recently, NEPA and other environmental laws have been challenged by the military, which has asserted that these laws interfere with combat readiness and, therefore, national security.[3] In *Winter v. National Resources Defense Council, Inc.*,[4] the Court invalidated portions of a preliminary injunction restricting the Navy's use of high-intensity sonar, finding that the balance of equities and public policy interests favored the Navy's interest in conducting realistic training exercises.[5] The Court's evaluation of the balance of equities and public policy interests, however, was based on its complete deference to the Navy's factual determinations regarding crucial aspects of the case.[6]

This Comment argues that the Court's complete deference to the Navy's factual determinations unfairly tipped the balance of equities and public policy interests in favor of the Navy. This made it impossible for the Court to accurately evaluate the propriety of injunctive relief. While a measure of deference to the military's professional expertise is desirable, excessive deference undermines other key government objectives.[7] By so deferring to the Navy's factual determinations, the Court enabled the military's ongoing rollback of environmental protection.

Part I of this Comment explains NEPA's environmental impact statement ("EIS") requirement, then discusses the standards that courts have used when evaluating the propriety of injunctive relief. Part II summarizes the Court's opinion in *Winter*, including the facts, procedur-

---

1. National Environmental Policy Act (NEPA) of 1969, 42 U.S.C. § 4321 (2006).
2. *Id.* § 4331(a):
> The Congress, recognizing the profound impact of man's activity on the interrelations of all components of the natural environment . . . declares that it is the continuing policy of the Federal Government . . . to use all practicable means and measures . . . to create and maintain conditions under which man and nature can exist in productive harmony, and fulfill the social, economic, and other requirements of present and future generations of Americans.

3. Marcilynn A. Burke, *Green Peace? Protecting Our National Treasures While Providing for Our National Security*, 32 WM. & MARY ENVTL. L. & POL'Y REV. 803 *passim* (2008).
4. 129 S. Ct. 365 (2008).
5. *Id.* at 378.
6. *See id.*
7. *See* Burke, *supra* note 3, at 874 ("[W]hile [the Department of Defense] asserts its need for this statutory relief to fulfill its obligation to defend this nation, it may be simultaneously abandoning much of what makes this country worth defending.").

al history, and opinions. Part III explores three topics: (1) the Court's deference to the Navy's factual determinations; (2) the impact that *Winter* may have on future military compliance with NEPA; and (3) the *Winter* decision in the context of post-9/11 trends in military environmental compliance. This Comment concludes that the *Winter* Court unreasonably deferred to the Navy, that *Winter* threatens to undermine NEPA, and that *Winter* highlights the military's ongoing rollback of environmental legislation in the wake of 9/11.

I. BACKGROUND

A. The National Environmental Policy Act of 1969

NEPA requires that every federal agency consider the environmental consequences of a proposed course of action before acting.[8] To this end, NEPA requires that an agency submit a detailed EIS prior to undertaking any activity that "significantly affect[s] the quality of the human environment."[9] When the action would result in no significant impact on the environment, NEPA permits agencies to file an environmental assessment ("EA") in lieu of an EIS, or to rely on a categorical exclusion if one applies.[10]

Crucially, NEPA does not require a particular substantive result in any case. Nonetheless, in *Robertson v. Methow Valley Citizens Council*,[11] the Court noted that NEPA does contain "action-forcing" procedures.[12] First, a timely EIS focuses an agency's attention on the environmental consequences of a proposed project, ensuring that the agency is aware of those consequences before starting the project.[13] Second, a timely EIS assures the public that the agency has taken environmental consequences into account during the decision-making process.[14] NEPA's implementing regulations also require that the agency accept public input once a draft EIS is released.[15] Third, a timely EIS gives notice to the various other governmental bodies that might have to deal with the project's offsite or secondary consequences.[16]

8. 42 U.S.C. § 4331(a)-(b) (2006).
9. *Id.* § 4332(2)(C).
10. *See* 40 C.F.R. § 1508.9 (2009) (explaining the Council on Environmental Quality's definition of "Environmental Assessment"); *id.* § 1508.13 (explaining the Council on Environmental Quality's definition of "[f]inding of no significant impact"); *id.* § 1501.4(2) (explaining that categorically excluded activities do not normally require an EIS).
11. 490 U.S. 332 (1989).
12. *Id.* at 348-51.
13. *Id.* at 349.
14. *Id.*
15. 40 C.F.R. § 1503.1(a)(4) (2009).
16. *Id.* § 1503.1(a)(1).

B. The Test for Evaluating the Propriety of Injunctive Relief

Preliminary injunctions are remedies at equity, issued only where there is a risk of irremediable injury, and where legal remedies would be inadequate.[17] Courts have traditionally evaluated the propriety of injunctive relief by balancing the inconvenience to the non-moving party (should the injunction be imposed) against the risk of irreparable harm to the moving party.[18] Courts also give attention to the public consequences of ordering injunctive relief.[19] Until *Winter*, the Ninth Circuit applied a more flexible standard, and Ninth Circuit courts were able to tailor injunctive relief to fit "the necessities of the particular case."[20] Some have argued that this flexible standard unduly increases the availability of injunctive relief, which has traditionally only been imposed in extraordinary circumstances.[21]

II. *WINTER V. NATURAL RESOURCES DEFENSE COUNCIL*

A. Facts

For the last forty years, the Navy has used mid-frequency active ("MFA") sonar during training exercises off of the southern coast of California ("SOCAL").[22] The MFA sonar system works by emitting high-intensity sound into the ocean depths, and then analyzing the echoes to reveal underwater objects.[23] MFA sonar is currently the only available means of detecting the near-silent submarines deployed by potential military adversaries.[24]

However, a "rapidly accumulating body of evidence" shows that MFA sonar injures marine mammals, both directly and indirectly.[25] The SOCAL waters are home to thirty-seven species of marine mammals.[26] Directly, MFA sonar generates energy sufficient to cause cranial hemorrhaging or decompression sickness in these animals.[27] Indirectly, MFA sonar harms marine mammals by contributing to underwater noise pollution.[28] Because many marine mammals have evolved to depend on sound, any significant increase in underwater noise pollution, such as the

17. Weinberger v. Romero-Barcelo, 456 U.S. 305, 311-12 (1982).
18. *Id.* at 312.
19. *Id.*
20. *Id.*
21. *See* Winter v. Natural Res. Def. Council, Inc., 129 S. Ct. 365, 375-76 (2008).
22. Brief for the Petitioners at 2, *Winter*, 129 S. Ct. 365 (No. 07-1239), 2008 WL 3285392.
23. *Winter*, 129 S. Ct. at 370.
24. *Id.* at 370-71.
25. *See* Joel R. Reynolds, *Submarines, Sonar, and the Death of Whales: Enforcing the Delicate Balance of Environmental Compliance and National Security in Military Training*, 32 WM. & MARY ENVTL. L. & POL'Y REV. 759, 760 (2008).
26. *Winter*, 129 S. Ct. at 371.
27. Reynolds, *supra* note 25, at 762-63.
28. *Id.* at 760.

use of MFA sonar, impairs marine mammals' breeding, hunting and navigation activities.[29]

B. Procedural History

In February 2007, the Navy released an EA for an upcoming series of SOCAL training exercises,[30] estimating that the exercises would result in over 500 severe and 170,000 minor injuries to marine mammals in SOCAL waters.[31] Nevertheless, the Navy concluded that this did not amount to the requisite "significant impact" on the environment, and therefore that a full EIS would not be necessary.[32]

Plaintiff National Resources Defense Council ("NRDC") filed a complaint against the Navy shortly thereafter, alleging that the Navy violated NEPA by proceeding with the exercises without first preparing an EIS.[33] NRDC sought to compel the Navy to prepare an EIS.[34] The district court found the Navy's EA inadequate, and that the EA did not relieve the Navy of its obligation to prepare a full EIS.[35] Accordingly, the district court enjoined the Navy's use of MFA sonar until the Navy prepared an EIS.[36] On appeal, the Ninth Circuit held that the district court's injunction was too broad, and remanded the case with instructions to impose mitigation provisions that would allow the Navy to continue its exercises.[37]

On remand, the district court imposed on the Navy a synthesis of six mitigation provisions.[38] The Navy challenged two of the six provisions on appeal: the mandatory shutdown of MFA sonar when a marine mammal is sighted within 2200 yards, and the mandatory seventy-five percent reduction in sonar volume during "surface ducting"[39] conditions.[40] While waiting for the appeal, the Navy sought emergency[41] Executive Branch relief from the Council on Environmental Quality ("CEQ"),[42] which au-

29. *Id.*
30. *Winter*, 129 S. Ct. at 372.
31. *Id.* at 392 (Ginsburg, J., dissenting).
32. *Id.* at 372 (majority opinion).
33. Winter v. Natural Res. Def. Council, Inc., 530 F. Supp. 2d 1110, 1113-14 (C.D. Cal. 2008).
34. Winter v. Natural Res. Def. Council, Inc., 518 F.3d 658, 661 (9th Cir. 2008).
35. *Winter*, 530 F. Supp. at 1116-17.
36. *Id.* at 1115.
37. *Winter*, 129 S. Ct. at 373.
38. *Winter*, 530 F. Supp. 2d. at 1118-21.
39. Surface ducting conditions "occur[] when the presence of layers of water of different temperature make it unusually difficult for sonar operators to determine whether a diesel submarine is hiding below." *Winter*, 129 S. Ct. at 385 (Breyer, J., concurring).
40. *Id.* at 373 (majority opinion). Hereinafter these two provisions will be referred to as "shutdown" and "power-down" provisions.
41. Under NEPA's implementing regulations, agencies may, in emergency circumstances, request the Executive Branch to authorize alternative arrangements that are limited to the scope of the emergency. 40 C.F.R. § 1506.11 (2009).
42. The Council on Environmental Quality is the Executive Branch office that is responsible for overseeing federal compliance with NEPA. 42 U.S.C. §§ 4341-47 (2006).

thorized the Navy to implement alternative arrangements of its own making.[43]

Alternative arrangements in hand, the Navy asked the Ninth Circuit to vacate the shutdown and power-down provisions.[44] The Ninth Circuit disregarded the alternative arrangements, finding that emergency circumstances did not exist, and the court upheld the two challenged provisions in the district court's injunction.[45] The Navy petitioned for certiorari on the issue of whether the district court abused its discretion in imposing the preliminary injunction, and the Supreme Court granted the petition.[46]

C. Majority Opinion

Chief Justice Roberts wrote the opinion of the Court, with Justices Scalia, Kennedy, Thomas, and Alito joining.[47] The majority opinion discussed the propriety of the shutdown and power-down provisions in the district court's preliminary injunction.[48] As such, it generally avoided addressing whether NEPA required the Navy to prepare an EIS before starting the sonar exercises.[49] The Court found that the shutdown and power-down provisions in the district court's injunction were an abuse of discretion, because the district court failed to accord sufficient deference to the Navy's factual determinations when evaluating the balance of equities and public policy interests.[50]

First, the Court adjusted the standard the Ninth Circuit used when deciding whether a preliminary injunction would be proper. The Ninth Circuit standard, prior to *Winter*, allowed the court to order an injunction when there was a "possibility" of irreparable harm to the moving party.[51] The Court found this standard to be too lenient, and established that there must instead be a "likelihood" of irreparable harm to the moving party.[52]

In light of the confusion surrounding this issue, Chief Justice Roberts synthesized precedent to explain a four-part test for evaluating the propriety of injunctive relief: "A plaintiff seeking a preliminary injunction must establish (1) that he is likely to succeed on the merits, (2) that he is likely to suffer irreparable harm in the absence of preliminary relief,

43. *Winter*, 129 S. Ct. at 373.
44. *Id.* at 374.
45. *Id.*
46. Winter v. Natural Res. Def. Council, Inc., 128 S. Ct. 2964 (2008) (mem.).
47. *See Winter*, 129 S. Ct. at 371, 382, 387.
48. *Id.* at 374-81.
49. *Id.* at 381.
50. *See id.* at 382.
51. *Id.* at 375.
52. *Id.*

(3) that the balance of equities tips in his favor, and (4) that an injunction is in the public interest."[53]

Furthermore, the Court found that the lower courts had failed to reconsider the likelihood of harm to the moving party in light of the shutdown and power-down provisions alone.[54] The Court acknowledged, however, that these errors might ultimately have been harmless:

> It is not clear that articulating the incorrect standard affected the Ninth Circuit's analysis of irreparable harm. Although the court referred to the possibility standard, and cited Circuit precedent along the same lines, it affirmed the District Court's conclusion that plaintiffs had established a near certainty of irreparable harm.[55]

Second, the Court briefly addressed the merits of the case: whether the Navy's EA had satisfied its NEPA obligations, without filing an EIS. The Court found that the Navy had met its procedural obligations under NEPA for two reasons: because the Navy's activities had taken place for forty years, and because the Navy had taken a "hard look at environmental consequences" before launching the SOCAL training exercises.[56]

Third, the Court discussed the balance of equities. The Court emphasized that, because the case involved matters of national security, the Court should grant "'great deference to the professional judgment of military authorities concerning the relative importance of a particular military interest.'"[57] With this in mind, the Court accepted the Navy's statement that the shutdown and power-down provisions would degrade the value of the SOCAL training exercises.[58]

The Court found that the Navy's interest in realistic training exercises outweighed the public's ecological, scientific, and recreational interests in obtaining a preliminary injunction, and therefore the shutdown and power-down provisions were an abuse of discretion.[59] The Court found it important that NRDC sued to compel the Navy to prepare an EIS, not to enjoin the Navy from the use of MFA sonar.[60]

Fourth, the Court addressed the Ninth Circuit's treatment of the public interest element. Once again, the Court stressed the need for deference to the military's judgment regarding the impact of mitigation ef-

---

53. *Id.* at 374 (emphasis added).
54. *Id.* at 376.
55. *Id.* (internal quotation marks omitted).
56. *Id.* (quoting Robertson v. Methow Valley Citizens Council, 490 U.S. 332, 350 (1989)).
57. *Id.* at 377 (quoting Goldman v. Weinberger, 475 U.S. 503, 507 (1986)).
58. *See id.*
59. *Id.* at 377-78.
60. *Id.* ("Given that the ultimate legal claim is that the Navy must prepare an EIS, not that it must cease sonar training, there is no basis for enjoining such training in a manner credibly alleged to pose a serious threat to national security.").

forts on the efficacy of training exercises.[61] Specifically, the Court found that the lower courts had not sufficiently deferred to the Navy's judgment regarding the magnitude of the burden imposed by the injunction.[62] While the Ninth Circuit found that any unreasonable burden on the Navy could later be alleviated by permitting the Navy to request relief on an emergency basis, the Supreme Court disagreed, arguing that it would compromise national security to award the Navy relief only once emergency circumstances have materialized.[63]

D. Concurring Opinion

Justice Breyer wrote a concurring opinion, in which Justice Stevens joined in part. The concurring opinion found the shutdown and power-down provisions to be an abuse of discretion for five reasons. First, Justice Breyer found that there was no proven need for the two challenged provisions.[64] In other words, there was no strong evidence establishing the amount of marginal harm attributable to the two specific challenged provisions.[65]

Second, Justice Breyer deferred to Navy officials' assertions that a delay in completing the SOCAL training exercises would have serious ramifications for the Navy's combat readiness.[66] Third, the concurring opinion found it important that the lower courts did not explain why they rejected the Navy's affidavit-supported assessment of the two challenged provisions' impact on the efficacy of the exercises.[67] Fourth, Justice Breyer addressed the Ninth Circuit's interpretation of the record, finding fault with the Ninth Circuit's conclusions.[68] Fifth, Justice Breyer argued that the district court imposed mitigation provisions that made continued exercises impossible, defying the Ninth Circuit's instructions.[69]

The concurrence concluded by suggesting that the Court should have modified the injunction to include the altered shutdown and power-down provisions that the Ninth Circuit imposed pending appeal.[70] Justice Breyer argued that the provisional measures had "become the status quo," and that the measures were equitable in light of the Navy's ability to conduct its training exercises under them.[71]

---

61. *See id.* at 378-79.
62. *Id.*
63. *Id.* at 380-81.
64. *Id.* at 383 (Breyer, J., concurring).
65. *Id.* at 383-84.
66. *Id.* at 384 ("Taken by themselves, [the Navy's] affidavits make a strong case for the proposition that insistence upon the two additional mitigating conditions would seriously interfere with necessary defense training.").
67. *Id.* at 384-85.
68. *Id.* at 385-86.
69. *See id.* at 386.
70. *Id.* at 387.
71. *Id.*

E. Dissenting Opinion

Justice Ginsburg wrote a dissenting opinion, in which Justice Souter joined. The dissent argued that the lower courts did not abuse their discretion, because the balance of equities favored upholding the preliminary injunction against the Navy.[72]

First, Justice Ginsburg discussed the case and determined that the plaintiffs were likely to succeed on the merits, in that an EIS was probably required before launching the SOCAL exercises.[73] She argued that the EIS is more than a mere procedural device; it forces government agencies to fully consider the impact of their actions on the environment before they act.[74] As such, failure to prepare a timely EIS defeats NEPA's purpose.[75] She emphasized that, while NEPA does not require a certain result, NEPA's EIS requirement plays a crucial informational role in the decision-making process.[76]

Justice Ginsburg argued that the Navy could have avoided this suit by simply preparing a sufficient, timely EIS.[77] The Navy sought favorable relief from the Executive Branch, rather than through the proper legislative channels.[78] Justice Ginsburg argued that this was "surely not what Congress had in mind when it instructed agencies to comply with NEPA 'to the fullest extent possible.'"[79] Therefore, because the Navy attempted to shortcut the process, equity demanded that the Navy bear the burden of its procedural failures in the form of a preliminary injunction.[80]

Second, Justice Ginsburg stressed the importance of flexibility in equitable relief.[81] When the likelihood of harm is high, the likelihood that the moving party will succeed on the merits is less important, and vice versa.[82] The dissent argued that the majority opinion does not reject this "sliding scale" standard.[83] Justice Ginsburg found it significant that the Navy itself predicted harm to marine mammals.[84] In light of this discussion, the dissent would have upheld the preliminary injunction in its entirety.[85]

---

72. *Id.* at 387 (Ginsburg, J., dissenting).
73. *Id.* at 393.
74. *See id.* at 389-90.
75. *Id.* at 390.
76. *Id.* at 389-90.
77. *See id.* at 390.
78. *Id.*
79. *Id.* at 391 (quoting 42 U.S.C. § 4332 (2006)).
80. *See id.* at 391-93.
81. *Id.* at 391 ("Flexibility is a hallmark of equity jurisdiction.").
82. *Id.* at 392.
83. *Id.*
84. *Id.*
85. *Id.* at 393.

III. ANALYSIS

In *Winter*, the Court weighed in on a pre-existing conflict between "the safety and continuation of the Republic and other values we hold dear, among them a healthy environment."[86] The Court correctly pointed out that the judiciary is inexperienced with national security issues, and therefore the Court should defer to the military's judgment regarding national security.[87] However, the Court's complete deference to the Navy unfairly tipped the balance of equities and public policy interests in favor of the Navy. More seriously, the Court's tacit approval of the Navy's actions in circumventing NEPA could effectively invalidate NEPA as it concerns the military. Finally, the Court's decision in *Winter* highlights a broader trend: the military has used 9/11 and national security concerns as a pretext for rolling back constraining environmental legislation.

A. The Court's Complete Deference to the Navy's Factual Determinations

The Court deferred extensively to the Navy regarding several of the Navy's key factual determinations. This deference, in aggregate, had the effect of skewing the balance of equities and public policy interests in favor of the Navy. The skewed balance of equities and public policy interests made it impossible for the Court to accurately evaluate the propriety of injunctive relief.

The Court's deference to the Navy in *Winter* involved the Navy's factual determinations, rather than the Navy's interpretations of its own enabling act or regulations. Therefore, the *Chevron*[88] and *Skidmore*[89] deference doctrines do not apply. Instead, the Court applied the military deference standard set out in *Goldman v. Weinberger*.[90] Professor Jonathan Masur assails this type of deference as "judicial abdication,"[91] and contends that it is inconsistent with the body of law requiring some level of judicial inquiry into agency determinations.[92] Professor Masur asserts that such deference to the military has "overwhelmed the legal strictures established to constrain the operation of executive power."[93] The Court's

86. Hope Babcock, *National Security and Environmental Laws: A Clear and Present Danger?*, 25 VA. ENVTL. L.J. 105, 107 (2007).

87. *Winter*, 129 S. Ct. at 377.

88. *Chevron* deference dictates that an agency's interpretation of its enabling act is binding on courts, so long as the enabling act's language is ambiguous and the interpretation of that language is reasonable. WILLIAM F. FOX, UNDERSTANDING ADMINISTRATIVE LAW 314-15 (5th ed. 2008) (discussing Chevron U.S.A., Inc. v. Natural Res. Def. Council, Inc., 467 U.S. 837 (1984)).

89. *Skidmore* deference dictates that an agency's interpretation of its own regulations is persuasive but not necessarily controlling. *Id.* at 320 (discussing Skidmore v. Swift, 323 U.S. 134 (1944)).

90. 475 U.S. 503, 507 (1986) ("[C]ourts must give great deference to the professional judgment of military authorities concerning the relative importance of a particular military interest.").

91. Jonathan Masur, *A Hard Look or a Blind Eye: Administrative Law and Military Deference*, 56 HASTINGS L.J. 441, 515 (2005).

92. *Id.* at 518-19.

93. *Id.* at 445.

complete deference to the military's factual determinations in *Winter* permitted the Navy's version of the facts to determine the balance of equities and public policy interests. The Court's opinion demonstrates the effect of this deference in five different ways.

First, the Court deferred to the Navy's claim that no evidence connected the forty years of SOCAL exercises with a single sonar-related injury to a marine mammal.[94] Yet, the Navy itself admitted that the exercises would affect approximately 80,000 marine mammals, some of which would be severely injured or killed.[95] In fact, in 2000, the Navy and NOAA Fisheries conducted an investigation into a mass marine mammal stranding event in the Bahamas.[96] The report concluded that the seventeen marine mammals were driven onto shore by injuries from underwater acoustic sources.[97] The report connected those injuries to a series of contemporaneous Navy MFA sonar exercises, and the Navy pledged to be more careful in the future.[98]

The evidence that the use of MFA sonar causes mass marine mammal strandings and deaths is "overwhelming," and the Navy was well aware of it.[99] It is surprising, then, that the Court deferred to the Navy's assertion that there would be no irremediable damage to the environment. It is difficult to think of an injury less remediable than the death of any number of marine mammals.

By contrast, the Navy's probable injuries in the case of a mid-training sonar shutdown are quite remediable. A mid-exercise MFA sonar shutdown would delay the completion of the exercise, and would undoubtedly raise costs, but it would not make completion of the exercise impossible.[100] The Navy mischaracterized this inconvenience as an irremediable injury, and the effect on marine mammals as negligible. The majority accepted this mischaracterization at face value.

Second, the Court observed that the injunction's shutdown provision would amount to a hundredfold increase in the surface area of the shutdown zone.[101] However, at the Navy's urging, the Court disregarded the observation that this MFA sonar shutdown zone is roughly the same size as the Navy's existing long-frequency active ("LFA") sonar shutdown zone.[102] The Court, perhaps humbled by the Navy's chastisement

94. Winter v. Natural Res. Def. Council, Inc., 129 S. Ct. 365, 381 (2008).
95. Brief for the Petitioners, *supra* note 22, at 9.
96. NAT'L OCEANIC & ATMOSPHERIC ADMIN. & DEP'T OF THE NAVY, JOINT INTERIM REPORT: BAHAMAS MARINE MAMMAL STRANDING EVENT OF 15-16 MARCH 2000, at ii (2001), www.nmfs.noaa.gov/pr/pdfs/health/stranding_bahamas2000.pdf.
97. *Id.*
98. *Id.* at vi.
99. *See* Reynolds, *supra* note 25, at 762-69 (providing a catalog of other sonar-related stranding events).
100. *See Winter*, 129 S. Ct. at 377; *id.* 384 (Breyer, J., concurring).
101. *Id.* at 379.
102. *See id.* at 379-80.

of the Ninth Circuit, declined to explore the effect on the training exercises of congruent MFA/LFA shutdown zones.[103] By deferring to the Navy's unsubstantiated claim that MFA sonar and LFA sonar are irreconcilably dissimilar in terms of the effect of the technology on marine mammals,[104] the Court failed to consider a range of factors that could have shown the burden to be smaller than the Navy asserted it to be.

Third, the Court deferred to the Navy regarding the power-down provision. The Court correctly recognized the Navy's important interest in training under surface ducting conditions when they exist.[105] Presumably, however, the conditions that conceal enemy submarines also conceal marine mammals. In other words, when surface ducting conditions exist, the Navy must be just as vigilant in avoiding marine mammals as it is in looking for enemy submarines. As Justice Breyer argued, the Court could have imposed the Ninth Circuit's provisional injunction, requiring the Navy to power down the sonar in proportion to the proximity of marine mammals to the vessel.[106] Justice Breyer's compromise would allow the Navy to continue training, while mitigating the injury to nearby marine mammals.

Fourth, the Court deferred to the Navy regarding the connection between the SOCAL training exercises and national security. The Navy asserted that the injunctions would jeopardize national security.[107] This conclusion was an exaggeration. The injunctions issued by the district court would not make training exercises impossible; they would merely cause delay and disruption.[108] Also, the injunctions applied to training exercises in SOCAL waters, and not to Navy actions generally.[109]

The Navy also argued the injunction would create "an unacceptable risk to the Navy's ability to train for essential overseas operations at a time when the United States is engaged in war in two countries."[110] This assertion was also an exaggeration. While the United States was indeed at war in Iraq and in Afghanistan, none of the United States' adversaries in those countries fielded a naval force—let alone the advanced "silent submarines" that MFA sonar was designed to detect. The Navy failed to explain the connection between adequate sonar training and combat readiness against these land-based, non-state forces. The Navy failed to explain how a delay in sonar training presented an "unacceptable risk" to

---

103. "The Ninth Circuit's willingness to strike out on its own in this complicated military and scientific context underscores the degree to which it has failed to defer to the professional military judgment of the Nation's most senior naval officers . . . ." Brief for the Petitioners, *supra* note 22, at 53.

104. *Winter*, 129 S. Ct. at 380.

105. *Id.*

106. *Id.* at 387 (Breyer, J., concurring).

107. *Id.* at 381 (majority opinion).

108. *Id.* at 384 (Breyer, J., concurring).

109. *See id.* at 372-73 (majority opinion).

110. Brief for the Petitioners, *supra* note 22, at 20.

ground forces in Iraq and Afghanistan.[111] The Navy also failed to explain how the injunction affected the combat readiness of already-deployed forces, other than underlining the importance of fleet-wide integration.[112] Professor Burke refers to such unsubstantiated claims as "thought-terminating cliché[s]."[113]

Finally, the Court deferred to the Navy regarding the urgent need to dissolve the injunction. The Court vacated the two challenged provisions, effectively handing the Navy the same result that it had sought from the CEQ.[114] As Justice Ginsburg points out, however, the "emergency circumstances" under which the Navy obtained the alternative agreements were of the Navy's own making: had the Navy filed an EIS before launching the SOCAL training exercises, it would not have had to seek emergency relief from the CEQ.[115] Justice Ginsburg correctly argued that the Navy should bear the burden of its procedural failures.[116]

The majority's eagerness to defer to the Navy's factual determinations regarding significant aspects of this case could also stem from the suit's underlying subject matter. The NRDC sought not to permanently enjoin the Navy from use of MFA sonar, but to compel the Navy to prepare an EIS, which at this stage of the case would have been merely a procedural gesture.[117] Nevertheless, the Court's complete deference to the Navy's factual determinations prevented the Court from accurately evaluating the propriety of injunctive relief.

B. Winter*'s Impact on Future Military Compliance with NEPA*

The underlying issue in *Winter* was whether the Navy should have prepared an EIS before launching its SOCAL training exercises. The majority opinion focuses on the narrower issue: the propriety of two challenged provisions in the district court's preliminary injunction. Although the Court generally avoided discussing the merits of the case, during the injunctive relief discussion it did briefly discuss the likelihood that NRDC would prevail on its NEPA claim.[118] The Court characterized NEPA as a mere procedural device, and implied that the Navy's EA satisfied NEPA's EIS requirement.[119]

---

111. *See id.* at 46 (stating that soldiers involved in the SOCAL training exercises conduct missions within Iraq and Afghanistan).

112. *Id.* at 4.

113. Burke, *supra* note 3, at 808.

114. *Winter*, 129 S. Ct. at 373-74.

115. *Id.* at 390 (Ginsburg, J., dissenting).

116. *See id.* at 387.

117. *See id.* at 381 (majority opinion).

118. *Id.* at 382.

119. *Id.* at 376 ("NEPA imposes only procedural requirements to 'ensur[e] that the agency, in reaching its decision, will have available, and will carefully consider, detailed information concerning significant environmental impacts.'" (alteration in original) (quoting Robertson v. Methow Valley Citizens Council, 490 U.S. 332, 349 (1989))).

The assertion that NEPA is a procedural device is technically true in that NEPA does not require any particular result,[120] but characterizing NEPA as mere procedure ignores the crucial role that NEPA procedures are designed to play in government agencies' decision-making processes. As Justice Stevens noted in *Robertson*, NEPA's EIS requirement ensures that government agencies will make informed decisions regarding whether to undertake a proposed action.[121] In other words, the EIS is a tool to be used *before* the action, to determine *whether* the action will take place.

Justice Ginsburg agrees: "[T]he timing of an EIS is critical. . . . An EIS must be prepared 'early enough so that it can serve practically as an important contribution to the decisionmaking process and will not be used to rationalize or justify decisions already made.'"[122] Justice Ginsburg characterized the Navy's actions as "short circuit[ing]" its NEPA obligations.[123]

Chief Justice Roberts' majority opinion argued that NEPA's ultimate objective is to ensure that agencies take a "hard look" at environmental consequences.[124] The Court argued that the Navy had conducted its SOCAL training exercises for forty years without incident, and the Navy's EA amounted to a hard look at the training exercise's environmental consequences.[125] Therefore, the Court concluded, the Navy satisfied this objective.[126]

The majority's reasoning is troubling. First, the Court's assertion that the use of MFA sonar should not be enjoined because it has been in use for forty years[127] is flawed. It would be bad public policy for a practice's history to establish its immunity from injunction. Otherwise, government agencies would have carte blanche to continue any established practice, regardless of harmful environmental impacts.[128]

Second, the Court construes NEPA to require merely that agencies take a hard look at environmental consequences.[129] However, NEPA unambiguously requires government agencies to prepare an EIS for all

120. *Robertson*, 490 U.S. at 350.
121. *Id.* at 349.
122. *Winter*, 129 S. Ct. at 390 (Ginsburg, J., dissenting) (quoting Andrus v. Sierra Club, 442 U.S. 347, 351 n.3 (1979)).
123. *Id.* at 391.
124. *Id.* at 376 (majority opinion) (quoting *Robertson*, 490 U.S. at 350).
125. *Id.*
126. *Id.*
127. *Id.*
128. The Court struck down another long-standing practice during the same term it decided *Winter*: In *Arizona v. Gant*, a Fourth Amendment case involving police officers' right to search an arrestee's vehicle, the court stated that, "[i]f it is clear that a practice is unlawful, individuals' interest in its discontinuance clearly outweighs any law enforcement 'entitlement' to its persistence." Arizona v. Gant, 129 S. Ct. 1710, 1723 (2009).
129. *Winter*, 129 S. Ct. at 376.

major federal actions significantly affecting the human environment.[130] Justices Ginsburg and Stevens have both discussed the importance of preparing an EIS, as opposed to merely taking a hard look.[131] Justices Ginsburg and Stevens have concluded that NEPA's clear procedures, including the EIS, are necessary to effectuate NEPA's purpose.[132]

The Court's decision suggests that NEPA compliance is arbitrary, so long as agencies take a hard look at the environmental consequences of an activity. If adopted in a future decision, this "hard look" standard would undermine NEPA by permitting agencies to provide post-hoc justification for environmentally destructive activities. Moreover, compared to the clear EIS guidelines,[133] this hard look standard is vague. The standard threatens to waste judicial resources by delegating to courts the responsibility to determine, on a fact-intensive case-by-case basis, what constitutes a hard look at an activity's environmental consequences. By contrast, if an agency prepares an EIS, all a reviewing court must do is apply NEPA's EIS requirements and implementing regulations.

Professor Hope Babcock highlights the fact that, while Congress provided for military waiver procedures under many environmental statutes, it did not provide for a waiver procedure to NEPA's EIS requirement.[134] Despite this omission, NEPA allows the military some flexibility under the Administrative Procedures Act ("APA"). The APA definition of "agency" excludes the exercise of military authority "in the field in time of war," and only agencies as defined by the APA are required to prepare an EIS.[135] Congress could have provided a waiver to the EIS requirement, but it did not, possibly because the Navy was already exempt from preparing an EIS for wartime field operations. Congress did not exempt the Navy from its obligation to prepare an EIS for domestic training exercises, such as those in SOCAL.

Justice Ginsburg agrees that the EIS requirement is central to NEPA's objectives.[136] She reasoned that the Navy's failure to file an EIS instigated the action, and therefore that the Court should have upheld the injunction.[137] This suggested holding is in perfect harmony with envi-

130. 42 U.S.C. § 4332(2)(C) (2006).

131. *See Winter*, 129 S. Ct. at 389-91 (Ginsburg, J., dissenting); Robertson v. Methow Valley Citizens Council, 490 U.S. 332, 348-53 (1989).

132. *See Winter*, 129 S. Ct. at 389 (Ginsburg, J., dissenting); *Robertson*, 490 U.S. at 350.

133. § 4332(2)(C).

134. The Clean Water Act, Clean Air Act, Toxic Substances Control Act, CERCLA, Safe Drinking Water Act, Resource Conservation and Recovery Act, Endangered Species Act, Marine Mammal Protection Act, Coastal Zone Management Act, and National Historic Preservation Act all contain provisions which allow an executive official (usually the President) to waive the military's obligation to comply with these acts in emergency circumstances. Babcock, *supra* note 86, at 110-16.

135. *See* 5 U.S.C. § 701(b)(1)(G) (2006); § 4332(2)(C).

136. *Winter*, 129 S. Ct. at 389 (Ginsburg, J., dissenting) ("The EIS is NEPA's core requirement.").

137. *Id.* at 387, 390, 393.

ronmental compliance considerations; if the Court showed an inclination to impose preliminary injunctive relief when an agency fails to file an EIS, agencies would be more likely to comply with NEPA. Unfortunately, the Court's decision in *Winter* may prove to discourage agencies from filing NEPA-compliant environmental impact statements.

C. The Military's Post-9/11 Offensive Against Environmental Legislation

Finally, the *Winter* Court's willingness to defer to the Navy's judgment and to allow the Navy to bypass clear NEPA requirements is part of a broader, more troubling trend. Professor Babcock accuses the Department of Defense ("DOD") of manipulating post-9/11 national security concerns to stage an offensive against constraining environmental legislation.[138]

Professor Babcock explains this trend in light of the broader post-9/11 erosion of civil liberties exemplified by the USA PATRIOT Act.[139] The USA PATRIOT Act, enacted in the months immediately following 9/11, was intended to enhance the government's power to combat terrorist threats, but had the additional effect of eroding civil liberties.[140] Until recently, the military had to resort to various statutory waiver systems to circumvent environmental legislation.[141] But military efforts to curtail environmental legislation found new traction in the post-9/11 and post-USA PATRIOT Act reality.[142]

For example, in the years immediately following the 9/11 terrorist attacks, the DOD convinced Congress to exempt the military from key areas of the Migratory Bird Treaty Act ("MBTA"), the Marine Mammal Protection Act ("MMPA"), and the Endangered Species Act ("ESA").[143] These exemptions were characterized as essential to national security.[144] This trend shows no sign of slowing.[145] In fact, the Navy urged the Court in *Winters* to view the Navy's MMPA exemption as evidence that other environmental regimes should necessarily be subordinated to military

---

138. "[T]he military is using the 'war on terrorism' as a Trojan horse to get out from under thirty years of constraining environmental laws it has never fully accepted." Babcock, *supra* note 86, at 110. "[T]he 9/11 attacks provided DOD with an opportunity that it seized to get relief from laws that it has resisted for decades." *Id.* at 153.

139. *See id.* at 120-26.

140. USA PATRIOT Act of 2001, Pub. L. No. 107-56, 115 Stat. 272.

141. Babcock, *supra* note 86, at 110-16 (describing the waiver processes available to the military under various environmental statutes).

142. Burke, *supra* note 3, at 811.

143. *See id.* at 808; Babcock, *supra* note 86, at 127-30.

144. U.S. Representative Bob Barr, R-GA, advocated for the continued abrogation of environmental legislation, characterizing the debate as a false choice between soldiers "surviv[ing] on the battlefield" and "trampling blades of grass." Burke, *supra* note 3, at 807.

145. Professor Babcock notes that the military has also set its sights on securing exemptions from the Resource Conservation and Recovery Act ("RCRA"), the Comprehensive Environmental Response, Compensation, and Liability Act ("CERCLA"), and the Clean Air Act ("CAA"). Babcock, *supra* note 86, at 132.

training.[146] Then-Vice President Dick Cheney referred to the post-9/11 restrictions on civil liberties as "the new normalcy."[147] These assertions suggest an intent to roll back all constraining environmental legislation, not just MMPA or NEPA, which should have given the Court pause.

With *Winter*, this troubling trend has spread to NEPA. The Court accepted the Navy's tenuous assertion that the SOCAL training exercises are necessary to ensure military preparedness.[148] Such deference to the Navy's factual determinations, and willingness to create military exemptions to existing environmental regimes, allows the military to dodge its environmental obligations.

CONCLUSION

While deference to the military's professional judgment is to a certain extent desirable, it is possible for courts to defer to an unreasonable extent. When a court unquestioningly accepts one party's characterization of a case, the court simply cannot accurately evaluate the propriety of injunctive relief. In *Winter*, the Court's complete deference to the Navy's factual determinations unfairly tipped the balance of equities and public policy interests against the plaintiffs.

The Court's complete deference to the Navy will likely have an impact far beyond the parties involved. First, the Court's decision implies that the military can comply with NEPA's objectives without having to comply with NEPA procedures. Second, the Court's decision perpetuates the military's offensive against "constraining" environmental legislation.[149] In *Winter*, the Court missed out on an opportunity to slow this trend, and prevent the military's rollback of environmental legislation.

*Ian K. London**

146. Brief for the Petitioners, *supra* note 22, at 15.
147. Babcock, *supra* note 86, at 125.
148. *See* Winter v. Natural Res. Def. Council, Inc., 129 S. Ct. 365, 377 (2008).
149. Babcock, *supra* note 86, at 110.

* J.D. Candidate, 2011. I would like to thank Professors David Thomson, Kevin Lynch, and Federico Cheever at the University of Denver Sturm College of Law for their invaluable input. I would also like to thank the outstanding Denver University Law Review board and editorial staff for their tireless work. Finally, I would like to thank Barry Zuckercorn, Robert Loblaw, and Wayne Jarvis for reminding me to always find humor in the law.